Revisited Relativistic Dirac-Hartree-Fock X-ray Scattering Factors for Neutral Atoms with $Z = 2 – 118$, and Chemically-Relevant Ions: All Cations, Selected Monovalent Anions, and the Excited (Valence) States of Carbon & Silicon.

by

Olukayode Shiroye

A Dissertation Submitted in Partial Fulfilment of the Requirements for

the Degree of Doctor of Philosophy in Computational Science

Middle Tennessee State University

October 25, 2022

Dissertation Committee:

Dr. Anatoliy Volkov, Chair

Dr. Abdul Khaliq

Dr. Cen Li

Dr. Preston MacDougall

DEDICATION

I dedicate this to my mum for allowing me to live her dream

&

My wife and kids for all their sacrifice towards this PhD degree

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr Anatoliy Volkov for his amazing support throughout this study. I also thank all my committee members: Dr Abdul Khaliq, Dr Cen Li and Dr Preston McDougall. I appreciate Dr John Wallin, Computational & Data Science Program MTSU, Dr Andrienne Friedli and Chemistry Department MTSU

ABSTRACT

This is a two-article dissertation structure which comprises of four chapters. Chapter one gives the general overview of the research interest, motivation and identified gaps in the literature. The chapter two and three compose of the first and second papers submitted for publication respectively. Chapter four discussed the conclusion drawn on the results obtained from the research and how this result compares to previous studies on X-ray scattering factor calculation.

For this dissertation, two papers have been submitted for publications on revisited relativistic Dirac-Hartree-Fock X-ray scattering factor for neutral atoms with Z = 2 – 118 (He – Og) and Chemically-Relevant Ions: All Cations, Selected Monovalent Anions, and the Excited (Valence) States of Carbon & Silicon. The X-ray scattering factor calculation used the recently developed DBSR_HF program [Zatsarinny & Froese Fischer (2016). *Comput. Phys. Comm.* **202**, 287 – 303] to calculate the fully relativistic Dirac-Hartree-Fock ground-state wavefunctions for all atoms with Z = 2 – 118 (He – Og) and 318 chemically relevant ions. The calculations use the extended average level scheme and include both the Breit interaction correction to the electronic motion due to magnetic and retardation effects, and the Fermi distribution function for the description of the nuclear charge density.

Using the total and orbital (spinor) energies, charge density maxima, atomic mean radii and means spherical radii (Guerra et al., 2017) for the neutral atoms and total electronic & ionization energies for ions, adequate comparison have been made between the results obtained and several previous studies. A newly developed Fortran program SF was used for a precise integration of the X-ray scattering factors by employing the DBSR_HF's B-spline representation of the relativistic one-electron orbitals. Interpolation of the obtained X-ray scattering factor has also been done in the

$0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} ranges using the recommended analytical functions for a four-term and five-term expansions.

The X-ray scattering factor values obtained from the uniform treatments of the all the species seem to represent an excellent compromise among all the previous studies and should be a good replacement for values in Volume C of the 2006 edition of International Table for Crystallography (Maslen, Fox & O'Keefe, 2006).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. INTRODUCTION[1]

1.1 Preface

X-ray diffraction is one of the most powerful and widely used experimental physicochemical techniques for studying the crystalline state of matter (Stout & Jensen, 1989; Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997; Ladd & Palmer, 2013). Indeed, most of the universities in the United States, including MTSU, and abroad operate at least one single crystal or a powder X-ray diffractometer, or both. The powder X-ray diffraction technique excels at qualitative and quantitative identification of crystalline phases in a sample, while single-crystal X-ray diffraction reveals identities and precise locations of atoms in a crystal structure with a further possibility to explore the chemical bonding (Stout & Jensen, 1989; Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997; Ladd & Palmer, 2013).

The X-ray diffraction method is based on the idea that the incident beam of X-rays interacts with and is scattered by the electronic shells of atoms in a crystal structure. The incoming X-rays cause electrons of atoms to oscillate which in turn produces a secondary (scattered) electromagnetic wave. X-ray scattering from a crystal is elastic which means that while the directions of the incident and scattered X-rays are different, the energy stays constant. The ability of an atom to scatter X-rays depends on the number and the distribution of its electrons – the heavier atoms with a greater number of electrons produce a stronger X-ray scattering signal. For a basic X-ray crystal structure analysis, it is safe to assume that each atom in a crystal structure occurs in its ground electronic state. If so, each atom type displays a characteristic scattering picture which is called the X-ray *scattering (form) factor*. The X-ray scattering factor is defined as a Fourier transform of the electron density of an atom (Stout

[1] Several parts of this introduction also appear in *Acta Crystallographica Section A: Foundations and Advances* (Olukayode, Froese Fischer & Volkov, 2022; 2023)

& Jensen, 1989; Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997; Ladd & Palmer, 2013):

$$f(\mathbf{S}) = \int \rho(\mathbf{r}) e^{2\pi \mathrm{i}\mathbf{S}\cdot\mathbf{r}} \mathrm{d}^3\mathbf{r} \quad (1)$$

where **r** is a vector in direct space, **S** is the so-called *scattering vector*, i is the imaginary unit ($\mathrm{i} = \sqrt{-1}$) and $\rho(\mathbf{r})$ is the electron density of an atom or a particular electronic shell. The scattering vector **S** has the length $S = |\mathbf{S}| = 2\ \sin\theta\ /\lambda$, where the angle θ is defined in such a way that 2θ is the angle between the directions of the incident and scattered (diffracted) X-rays (Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997), and λ is the wavelength of the incident X-ray beam. Switching to the spherical coordinate system, i.e. $\mathbf{r} \equiv (r, \theta, \phi)$ and $\rho(\mathbf{r}) = \rho(r, \theta, \phi)$, and assuming that **S** is directed along the *z*-axis of the Cartesian coordinate system, in which case

$$\mathbf{S}\cdot\mathbf{r} = S\, r \cos\theta = 2\ \sin(\theta)\ /\lambda\, r \cos\theta = 2\, s\, r \cos\theta \quad (2)$$

where $s = \sin\theta\ /\lambda = S/2$, the integral (1) can be re-written as

$$f(s) = \int_0^{\infty}\int_0^{2\pi}\int_0^{\pi} \rho(r, \theta, \phi) e^{4\pi \mathrm{i} s r \cos\theta} r^2 \sin\theta\ \mathrm{d}\theta\ \mathrm{d}\phi\ \mathrm{d}r \quad (3)$$

In Table 1, we show the X-ray scattering factors $f(s)$ at the selected $s = \sin\theta\ /\lambda$ points between 0 and 6 Å^{-1} for five species that have the same number of electrons but different nuclear charge (these are called “isoelectronic” species).

Table 1 X-ray scattering factors of the isoelectronic species F^-, Ne, Na^+, Mg^{2+} and Al^{3+} from the Rez, Rez & Grant (1994) study.

$\sin\theta/\lambda$ (Å^{-1})	F^-	Ne	Na^+	Mg^{2+}	Al^{3+}
0.00	10.0000	10.0000	10.0000	10.0000	10.0000
0.05	9.7338	9.8303	9.8832	9.9139	9.9336
0.10	9.0157	9.3519	9.5463	9.6625	9.7383
0.15	8.0324	8.6442	9.0262	9.2655	9.4257
0.20	6.9742	7.8062	8.3746	8.7523	9.0137
0.25	5.9712	6.9291	7.6475	8.1571	8.5239
0.30	5.0878	6.0809	6.8961	7.5149	7.9800
0.35	4.3430	5.3039	6.1614	6.8576	7.4053
0.40	3.7318	4.6187	5.4724	6.2115	6.8209
0.45	3.2389	4.0308	4.8464	5.5968	6.2450
0.50	2.8461	3.5365	4.2917	5.0268	5.6916
0.60	2.2910	2.7912	3.3965	4.0480	4.6905
0.70	1.9475	2.2964	2.7535	3.2890	3.8604
0.80	1.7312	1.9720	2.3058	2.7248	3.2040
0.90	1.5872	1.7572	1.9978	2.3160	2.7018
1.00	1.4825	1.6098	1.7852	2.0234	2.3260
1.20	1.3236	1.4180	1.5246	1.6621	1.8427
1.40	1.1856	1.2808	1.3673	1.4605	1.5740
1.60	1.0530	1.1586	1.2464	1.3265	1.4102
1.80	0.9262	1.0416	1.1371	1.2185	1.2931
2.00	0.8080	0.9292	1.0325	1.1198	1.1955
2.50	0.5625	0.6808	0.7909	0.8902	0.9776
3.00	0.3881	0.4896	0.5912	0.6892	0.7808
3.50	0.2695	0.3514	0.4380	0.5260	0.6128
4.00	0.1895	0.2540	0.3247	0.3998	0.4770
5.00	0.0994	0.1377	0.1827	0.2330	0.2890
6.00	0.0547	0.0787	0.1074	0.1409	0.1789

Table 1 illustrates the ability of the X-ray diffraction analysis to distinguish the various atom and ion types. Indeed, despite all five species having the same number of electrons, the scattering factors for $\sin\theta/\lambda > 0$ are different.

Because the X-ray scattering factors are usually calculated for a fixed set of $\sin\theta/\lambda$ grid points (see, for example, Table 1), it is convenient to introduce a simple interpolating function that allows one to evaluate the X-ray scattering factor at the arbitrary $\sin\theta/\lambda$ value, thus bypassing a time-consuming integration of the electron density. The conventional interpolations of the X-ray scattering factors (Doyle & Turner, 1968; Cromer & Mann, 1968b; Cromer & Waber, 1968; Thakkar & Smith, 1992; Maslen, Fox & O'Keefe, 2006) include the function

$$f(\sin\theta/\lambda) = \sum_{i=1}^{m} a_i \exp(-b_i \sin\theta^2/\lambda^2) + c \tag{4}$$

with $m = 4$ in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} interval as proposed by Vand, Eiland & Pepinsky (1957), and the function

$$f(\sin\theta/\lambda) = \exp\left(\sum_{i=0}^{n} a_i (\sin\theta/\lambda)^i\right) \tag{5}$$

with $n = 3$ for the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range as suggested by Fox, O'Keefe & Tabbernor (1989).

1.2 The importance of the electron density in the X-ray scattering factors calculation

The electron density is an important property of an atom or a molecule as it provides useful information about the reactivity of the species. These include the potential sites of nucleophilic and electrophilic attack, multipole moments of charge distribution, partial atomic electronic charge, electrostatic potential and electric field (Coppens, 1997).

Theoretically, the simplest way to obtain the electron density of an atom is through the atomic orbital wavefunctions which are the solutions of the non-relativistic many-electron time-independent electronic Schrödinger equation (Atkins & de Paula, 2006):

$$\hat{H}\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = E\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \tag{6}$$

$$\widehat{H} = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}\right) + \sum_{i<j}^{N}\frac{1}{r_{ij}} \tag{7}$$

where N is the number of electrons, Z is the nuclear charge, $\widehat{H}$ is the Hamiltonian operator (or simply "Hamiltonian"), ∇_i^2 is the kinetic energy operator for the i-th electron, $\mathbf{r}_i$ is the position of the i-th electron relative to the atomic nucleus, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $1/r_{ij}$ is the two-particle interaction operator, and $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$ is the N -electron wavefunction, that is, $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$ depends on the locations of all N electrons (Atkins & de Paula, 2006). In general, Ψ should also include the spin function of each electron but for now we ignore this dependence (Atkins & de Paula, 2006). This Schrödinger equation cannot be solved exactly because of the presence of the $1/r_{ij}$ term which describes the interelectronic (Coulomb) repulsion. Within the *Hartree-Fock approximation*, it is assumed that each electron is moving in an average field created by all other electrons (Atkins & de Paula, 2006). Thus, $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ is replaced by a set of wavefunctions $\{\psi\}$, each describing a behavior of one or two electrons, and each wavefunction is a function of the three spatial coordinates, $\mathbf{r} \equiv (x, y, z) \equiv (r, \theta, \phi)$, and a spin coordinate (α or β) (Atkins & de Paula, 2006). For example, the notation $\psi_1(\boldsymbol{r}_1, \alpha)$ and $\psi_1(\boldsymbol{r}_1, \beta)$ means that two electrons are described by the same spatial wavefunction $\psi_1(\boldsymbol{r}_1)$ but according to the *Pauli exclusion principle*, they must have different spins (α and β). These wavefunctions are called the *atomic orbitals* (AO), and the approximation is known as the *orbital approximation* (Atkins & de Paula, 2006). The advantage of the orbital approximation is that it replaces a single Schrödinger equation that includes $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ that describes all electrons in the system with N_{AO} Schrödinger equations (Atkins & de Paula, 2006):

$$\begin{aligned} \widehat{H}_1\psi_1(\boldsymbol{r}_1) &= E_1\psi_1(\boldsymbol{r}_1) \\ \widehat{H}_2\psi_2(\boldsymbol{r}_2) &= E_2\psi_2(\boldsymbol{r}_2) \end{aligned} \tag{8}$$

… … … … … … … … …

$$\widehat{H}_{N_{\mathrm{AO}}}\psi_{N_{\mathrm{AO}}}(\boldsymbol{r}_{N_{\mathrm{AO}}}) = E_{N_{\mathrm{AO}}}\psi_{N_{\mathrm{AO}}}(\boldsymbol{r}_{N_{\mathrm{AO}}})$$

where N_{AO} is the number of atomic orbitals, and E_i and $\widehat{H}_i$ are the energy and the Hamiltonian of the i-th atomic orbital, respectively. The latter is defined as (Atkins & de Paula, 2006):

$$\widehat{H}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} + \sum_{j\neq i}^{N}\frac{1}{r_{ij}} \tag{9}$$

Note that despite increasing the number of equations that we need to solve, each equation describes the behavior of only one or two electrons. In the case when the wavefunction ψ_i describes two electrons, the spatial part of the wavefunction is the same for both electrons, and only the spin part is different.

Using the non-relativistic quantum mechanical treatment of an atom, the electron density $\rho(\mathbf{r}) \equiv \rho(r,\theta,\phi)$ can be expressed in terms of the atomic orbital wavefunctions $\psi(\mathbf{r}) \equiv \psi(r,\theta,\phi)$

$$\rho(r,\theta,\phi) = \sum_{i=1}^{N_{\mathrm{occ}}} n_i|\psi_i(r,\theta,\phi)|^2 \tag{10}$$

where ψ_i and n_i are the wavefunction and the electron population (occupancy) of the i-th orbital, respectively, N_{occ} is the number of the occupied orbitals, and the square modulus $|x|^2$ is equal to x^2 if the quantity x is real and x^*x if it is complex (symbol * denotes the complex-conjugate). Representing wavefunction of the i-th atomic orbital, $\psi_i(r,\theta,\phi)$, as a product of the radial function $R_i(r)$, which is always real and normalized,

$$\int_0^\infty R_i^2(r)r^2\,\mathrm{d}r = 1 \tag{11}$$

, and the complex spherical harmonic angular function $Y_i(\theta,\phi)$ (Weissbluth, 1978), the expression for the electron density becomes

$$\rho(r,\theta,\phi) = \sum_{i=1}^{N_{\text{occ}}} n_i R_i^2(r)|Y_i(\theta,\phi)|^2 \tag{12}$$

For a spherically symmetric atom, which appears to be an excellent first-order approximation (Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997), the angular part is just

$$Y(\theta,\phi) = (4\pi)^{-1/2} \tag{13}$$

Introducing a simplified notation for the radial part of the non-relativistic atomic electron density

$$\mathbb{R}(r) = r^2 \sum_{i=1}^{N_{\text{occ}}} n_i R_i^2(r) \tag{14}$$

, the final expression for the atomic X-ray scattering factor becomes:

$$f(s) = \frac{1}{4\pi} \int_0^\infty \mathbb{R}(r) \frac{\sin(4\pi sr)}{sr} dr \tag{15}$$

There have been a number of studies (Clementi & Roetti, 1994, Stewart, 1976, Fischer, 1977, Hansen & Coppens, 1978, Bunge, Barrientos & Bunge, 1992, Bunge, Barrientos & Bunge, 1993, Koga, Tatewaki & Thakkar, 1993, Koga, Watanabe, Kanayama, Yasuda & Thakkar, 1995, Koga, Watanabe, Kanayama, & Thakkar, 1999) that used the non-relativistic Schrödinger equation to determine the atomic orbital wavefunctions from which the electron density and ultimately the atomic X-ray scattering factors could be calculated. However, it has been observed that accuracy of the X-ray scattering factors can be improved by including the relativistic effects which are particularly important for heavier atoms. Indeed, the inner electrons in heavy atoms are moving so fast that their speeds approach that of light (the speed of light).

1.3 Why the relativistic X-ray scattering factors

Relativity can have a profound effect on properties of heavy elements such as the lanthanides and actinides (Kaldor & Wilson, 2003). However, even for lighter elements, such

as those in the second period, the relativistic effects may be extremely important. Consider a neutral carbon atom with the nuclear charge $Z = 6$ in the ground-state electronic configuration $1s^2\ 2s^2\ 2p^2$, where 1s, 2s and 2p are the electronic subshells, each occupied by two electrons represented by the superscripts. The inner 1s shell is called the *core* shell, while the outer 2s and 2p subshells comprise a *valence* shell. Electrons occupying core shell(s) are called the *core electrons*, and those occupying valence shell(s) are called the *valence electrons*. While the two 1s electrons are the closest to the nucleus, they experience Coulomb repulsion not only from each other but also due to a small penetration of the 2s and 2p electrons into the core region. By the same token, the 1s electrons shield both the 2s and 2p electrons from the nucleus but this effect is much more significant. Finally, due to a deeper penetration into the core region, the 2s electrons are expected to experience a slightly higher nuclear charge as compared to the 2p electrons. The combined effect of the electron repulsion, penetration and shielding can be modeled by lowering the true nuclear charge, Z, to the *effective nuclear charge*, Z_{eff}, separately for each subshell:

$$Z_{\text{eff}} = Z - \sigma \tag{16}$$

where σ is known as the *shielding constant* of a subshell. The shielding constant σ is a positive quantity constrained such as $0 \leq \sigma < Z$. When $\sigma = 0$, the electrons are exposed to ("see" or "feel") the entire nuclear charge Z because $Z_{\text{eff}} = Z$, while when $\sigma \to Z$ and $Z_{\text{eff}} \to 0$, the electrons "feel" or "see" very little of the true nuclear charge. Both the non-relativistic (Clementi & Raimondi, 1963) and relativistic (Guerra et al., 2017) calculations of the effective nuclear charge parameter Z_{eff} for electrons in the three subshells of the carbon atom (Table 2) correctly predict the greatest Z_{eff} for the 1s electrons and the lowest Z_{eff} for the 2p electrons, though the relativistic calculations show a much more pronounced differences between Z_{eff} for the 2s and 2p subshells.

Table 2 The effective nuclear charge parameter Z_{eff} for electronic subshells in the ground-state carbon and silicon atoms from the non-relativistic (Clementi & Raimondi, 1963) and relativistic (Guerra et al., 2017) calculations.

Subshells	**Carbon atom** ($Z = 6$) $1s^2\ 2s^2\ 2p^2$		**Silicon atom** ($Z = 14$) $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$	
	Non-relativistic	Relativistic	Non-relativistic	Relativistic
1s	5.67	5.59	13.58	13.46
2s	3.22	3.78	**9.03**	**10.67**
2p	3.14	2.92	**9.95**	**9.34**
3s	—	—	4.90	6.13
3p	—	—	4.29	4.54

Now consider a silicon atom (Z=14) located in the third period of the periodic table just below carbon (Table 2). The non-relativistic calculations (Clementi & Raimondi, 1963) give Z_{eff} for the 2s subshell (9.03) which is actually lower (!) than that for the 2p shell (9.95). However, the relativistic calculations (Guerra et al., 2017) not only correctly predict the sequence of the Z_{eff} parameters for all subshells, but also show a more pronounced differences for Z_{eff} between the 3s and 3p subshells in comparison to their non-relativistic counterparts. In fact, the differences in Z_{eff} obtained from the non-relativistic and relativistic calculations start to show up as early in the third period as sodium ($Z = 11$). This example shows that the relativistic effects may be important even for lighter elements.

While not necessarily true, Autschbach (2012) argued that the effect of relativity on properties of some elements is not given sufficient attention apart from in spectroscopy and structural chemistry. Since the effect of gravity on very small mass particles such as proton, neutron and electron are infinitesimally negligible, gravitational effects on the chemical properties of atoms and molecules are insignificant (Autschbach, 2012). However, the effect

of relativity starts becoming apparent as the speed with which the electrons move tend to the speed of light. Relativistic effect is not due to the size of the screened nuclear charge, which is small for heavy atoms (Dirac, 1929), but due to possibly large kinetic and potential energies of the heavy atoms (Schwarz, 1989).

Relativity can significantly affect the chemical and physical properties of heavy elements and that is particularly obvious in the lower third elements of the periodic table (Autschbach, 2012). Even for elements that are light, factoring in the effect of relativity can produce a more accurate result as shown in the example above and as discussed, for instance, in the works of Kołos & Wolniewic (1964, 1965, 1968) and Michauk & Gauss (2007). As such, including the relativistic effect when calculating X-ray scattering factors should yield more accurate results.

In order to account for the relativistic effects, the Schrödinger equation is replaced with the Dirac-Hartee-Fock equation. For an N-electron atom or ion, the total Dirac-Coulomb Hamiltonian $\widehat{H}$ includes (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016)

$$\widehat{H} = \sum_{i=1}^{N} H_i\,(\mathbf{r}_i) + \sum_{i<j}^{N} \frac{1}{r_{ij}} \tag{17}$$

where $\mathbf{r}_i$ is the position of the i-th particle (electron) relative to the atomic nucleus, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, $1/r_{ij}$ is the two-particle interaction operator, and H_i is the one-electron Dirac operator for the i-th particle (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016):

$$H_i(\mathbf{r}_i) = c\boldsymbol{\alpha}.\,\mathbf{p}_i + \beta c^2 + V_{\mathrm{nuc}}(r_i) \tag{18}$$

where $\mathbf{p}_i$ is momentum operator of the i-th electron, c is the speed of light, $V_{\mathrm{nuc}}(r_i)$ is the nuclear potential and $\boldsymbol{\alpha}$ and β are Dirac matrices (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016). Within the *point model* of a nucleus of charge Z, $V_{\mathrm{nuc}}(r_i)$ is given by the Coulomb potential (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016)

$$V_{\text{nuc}}(r) = -Z/r \tag{19}$$

However, several more advanced models that take into account the finite size of a nucleus, such as the *uniform* or *Fermi charge distribution* (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016), can also be used. The total N-electron wavefunction is now represented by the anti-symmetrized products of Dirac four-component spinors (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016):

$$\Phi_{nk\mu}(\mathbf{r}) = \frac{1}{r}\begin{bmatrix} P_{nk}(r) & \chi_{k\mu}(\theta,\phi) \\ \mathrm{i}Q_{nk}(r) & \chi_{-k\mu}(\theta,\phi) \end{bmatrix} \tag{20}$$

where $\chi_{k\mu}(\theta,\phi)$ is the spin-angular function called the spinor spherical harmonic (Grant et al., 1980; Zatsarinny & Froese Fischer, 2016) which is analogous to the spherical harmonic function $Y(\theta,\phi)$ used in non-relativistic quantum mechanics. The radial density $\mathcal{R}(r)$ in relativistic calculation is formed from two relativistic subshells called the *major*, $P(r)$, and *minor*, $Q(r)$, components, and is subject to the following normalization condition:

$$\int_0^\infty [P^2(r) + Q^2(r)]\,\mathrm{d}r = 1 \tag{21}$$

1.4 The literature survey

Several studies have been published on the relativistic X-ray scattering factors for neutral atoms and chemically relevant ions. These include the works of Doyle & Turner (1968), Cromer & Waber (1968), Rez, Rez & Grant (1994), Wang, Smith, Bunge & Jáiregui (1996), Su & Coppens (1994, 1997, and 1998) and Macchi & Coppens (2001).

The original Dirac-Hartree-Fock atomic X-ray scattering factors were calculated for all neutral atoms with $Z = 2 - 98$ by Doyle & Turner (1968) based on wavefunctions of Coulthard (1967), and by Cromer & Waber (1968) using the wavefunctions of Mann (1968). The Coulthard (1967) approach included the point-charge model while Mann (1968) accounted for a finite charge distribution of the atomic nucleus, and both neglected the magnetic and

retardation effects. The Doyle & Turner (1968) calculations also included several cations and anions (Li^+, Be^{2+}, Na^+, Mg^{2+}, Cl^-, K^+, Ca^{2+}, V^{2+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^+, Zn^{2+}, Br^-, Rb^+, Sr^{2+}, Sn^{2+}, Sn^{4+}, I^-, Cs^+), while the rest of the chemically-relevant cations were calculated by Cromer & Mann (1968a) using the *non-relativistic* numerical Hartree-Fock wavefunctions of Mann (1967) and by Cromer & Waber (1968) using the *relativistic* Dirac-Slater method.

In 1994, Rez, Rez & Grant (1994, 1997) redetermined the X-ray scattering factors for naturally occurring elements and their selected ions using multiconfiguration Dirac-Hartree-Fock combined with the extended average level (EAL) model (Grant, Mayers & Pyper, 1976) as was implemented in the Oxford MCP/MCDF package (Grant et al., 1980) and later included in the GRASP suite of programs (Dyall et al., 1989). Regrettably, it is not known which nuclear model was used by Rez, Rez & Grant (1994) and whether the Breit interaction correction was applied. In addition, their work covered only a subset of ions calculated in the earlier studies (Doyle & Turner, 1968; Cromer & Waber, 1968; Cromer & Mann, 1968a) though they did include the X-ray scattering factors for Cr^{4+} and O^{2-} that had not been reported before. The calculations of Dirac-Hartree-Fock wavefunctions for anions (O^{2-}, F^-, Cl^-, Br^-, and I^-) in the Rez, Rez & Grant (1994) work required the use of the Watson sphere approximation (Watson, 1958) which involves surrounding of anions by a sphere of positive charge. The authors found "*very little difference (less than 0.1%) between* [their] *results and those of Doyle & Turner (1968) except for some heavy elements at high values of* $\sin\theta/\lambda$" despite the expected differences between their multiconfiguration DHF wavefunctions and single-configuration DHF wavefunctions used by Doyle & Turner (1968) and Cromer & Waber (1968). For interpolation, Rez, Rez & Grant (1994, 1997) used parametrization (39) with $m = 4$ but excluded the constant c. The two interpolating functions were designed to provide a high accuracy fit for $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and a lower accuracy fit for the entire 0 – 6 Å^{-1} range (Rez, Rez & Grant, 1994, 1997).

The 1996 calculations by Wang et al. (1996) were also performed at the multiconfiguration DHF level but were limited to neutral atoms with Z = 2 (He) – 18 (Ar), six ions (Li^+, Be^{2+}, $F^–$, Na^+, Mg^{2+} and $Cl^–$), and the excited (valence) states of carbon (C_{val}) and silicon (Si_{val}), though no additional approximations were included when calculating the two anions. An important feature of the study was that the relativistic wavefunctions approached the ground-state non-relativistic wavefunctions as the speed of light approached the infinity. The authors also chose to generate scattering factors on a fine grid instead of providing analytical interpolating functions.

Su & Coppens (1997) pointed out that the Rez, Rez & Grant (1994, 1997) wavefunctions may not correspond to the optimized ground states of the atoms and offered their own X-ray scattering factors (Su & Coppens, 1997, 1998) for neutral atoms with Z = 1 – 54 (Xe) computed using the multiconfiguration DHF method with the optimal level (OL) model in GRASP92 (Parpia, Froese Fischer & Grant, 1996). In 2001, Macchi & Coppens (2001) extended the work of Su & Coppens (1997, 1998) to all chemically relevant ions up to $I^–$. As in the work of Wang et al. (1996), no additional approximations were introduced when calculating the monovalent anions ($O^–$, $F^–$, $Cl^–$, $Br^–$, and $I^–$) but unlike the Rez, Rez & Grant (1994) study, the $O^{2–}$ and Cr^{4+} ions were not calculated. It is a little surprising that Volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) does not include the values from Su & Coppens (1997, 1998) and Macchi & Coppens (2001). Perhaps, it is related to the fact that Su they did not process all the atoms and ions listed in the currently used tables (Maslen, Fox & O'Keefe, 2006), and/or because the interpolating procedure for the scattering factors employed in their studies was somewhat different from the established approach: they used the function (39) with $m = 6$ and $c = 0$ in each of the 0 – 2, 2 – 4 and 4 – 6 $Å^{-1}$ $\sin\theta/\lambda$ intervals. Even after the passing of Prof. Coppens in 2017, the scattering factors from these studies are available online at

http://harker.chem.buffalo.edu/group/ptable.html. However, one should be aware that for several atoms some parts of the data are missing.

In the light of the above, there is clearly a need for a set of more uniformly calculated relativistic X-ray scattering factors for all neutral atoms and chemically relevant ions including those missing in the previous studies, which is the main goal of this work.

1.5 Using the B-spline Dirac-Hartree-Fock approach

The main goal outlined above will be reached by i) calculating the fully-relativistic wavefunctions of neutral atoms and ions using the B-spline Dirac-Hartree-Fock approach recently proposed and coded in the Fortran program DBSR_HF by Zatsarinny & Froese Fischer (2016), and ii) using the newly developed Fortran program SF (Olukayode, Fischer & Volkov, 2022) to integrate the X-ray scattering factors and to create the analytical interpolating functions.

B-spline is a piecewise polynomial function defined as follows:

$$B_{i,1}(r) = \begin{cases} 1, & t_i \le r \le t_{i+1} \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

$$B_{i,k}(r) = \frac{r - t_i}{t_{i+k-1} - t_i} B_{i,k-1}(r) - \frac{t_{i+k} - r}{t_{i+k} - t_{i+1}} B_{i+1,k-1}(r) \tag{23}$$

The functions are divided into intervals with end points t_i as the knots. Equation (23) is the de Boor algorithm (de Boor, 1971) for evaluation of the values of $B_{i,k}(r)$.

B-splines have been successfully used in both the non-relativistic (Froese Fischer, 2007; 2011) and relativistic (Johnson & Sapirstein, 1986) electronic structure calculations. One of the major advantages offered by a B-spline expansion is that the calculations are reduced to matrix algebra methods instead of the more complicated differential equations when using other basis sets (Zatsarinny & Froese Fischer, 2016). Perhaps, the only serious issue associated with using B-splines is the appearance of the pseudo-states when solving the Dirac equations though this problem is shared by all finite-element bases (Drake & Goldman, 1981; Zatsarinny

& Froese Fischer, 2016). In DBSR_HF, that problem is solved by using different B-spline orders for the small (kq) and large (kp) components with $kq = kp + 1$ (Froese Fischer & Zatsarinny, 2009). Then, the major and minor components of the radial function for the j-th relativistic atomic orbital (spinor) can be obtained from the B-spline expansion as follows

$$P_j(r) = \sum_{i=1}^{np} [p_{ji} B_i^{kp}(r)] \tag{24}$$

$$Q_j(r) = \sum_{i=1}^{nq} [q_{ji} B_i^{kq}(r)] \tag{25}$$

By default, the DBSR_HF code sets the values of kp and kq to 8 and 9, respectively. However, the B-spline bases for the major and minor components are formed on the same number of intervals. The p_{ji} and q_{ji} are orbital expansion coefficients and the np and nq are the grid points. The relativistic radial electron density $\mathcal{R}(r)$ can then be evaluated as

$$\mathcal{R}(r) = \sum_{j=1}^{N_{\text{occ}}} n_j \left[P_j(r)^2 + Q_j(r)^2 \right] \tag{26}$$

where n_j is the number of electrons of the j-th relativistic atomic orbital (spinor), and N_{occ} is number of the relativistic atomic orbitals. Then, the expression for the relativistic atomic or ionic X-ray scattering factor becomes essentially identical to its non-relativistic counterpart:

$$f(s) = \frac{1}{4\pi} \int_0^\infty \mathcal{R}(r) \frac{\sin(4\pi sr)}{sr} dr \tag{27}$$

1.6 The motivation, goal and objectives of the dissertation

There are two major reasons for revisiting the existing relativistic Dirac-Hartree-Fock X-ray scattering factors. The first is to produce the X-ray scattering factors for all atoms and chemically relevant ions that have not been previously published. The second motivation is to provide corrections for a number of errors and typos, and to improve the accuracy of the

existing X-ray scattering factors. We note that despite a number of more recent publications, the *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589 (Maslen, Fox & O'Keefe, 2006) still includes the X-ray scattering factors from 1968 calculations that were either performed at the non-relativistic level or replaced the Dirac-Hartree-Fock method with the Dirac-Slater approximation, or ignored important relativistic corrections to the electronic motion due to magnetic and retardation effects. Perhaps, the main reason for retaining the old data is that the newer calculations did not cover the same range of atoms and ions as the 1968 studies, and deviated from the established procedures for analytical interpolations.

As such, the overarching goal of the present work is to complete a uniform treatment of the X-ray scattering factors at the relativistic Dirac-Hartree-Fock level for

1) all neutral atoms with $Z = 2\ (\mathrm{He}) - 118\ (\mathrm{Og})$,
2) all chemically relevant cations of the elements with Z = 3 (Li) – 104 (Rf) (Greenwood & Earnshaw, 1997),
3) selected monovalent anions (O^-, F^-, Cl^-, Br^-, I^-, which were included in most of the earlier studies, plus At^-),
4) the ns^1np^3 excited (valence) states of carbon and silicon, and
5) several exotic cations (Db^{5+}, Sg^{6+}, Bh^{7+}, Hs^{8+} and Cn^{2+}),

thus, significantly extending the list of species that were treated in all the previous studies.

The objectives of this study are to

a) integrate the X-ray scattering factors with a high precision at all points of a fine $\sin\theta/\lambda$ grid between 0 and 6 $Å^{-1}$ (Wang et al., 1996);
b) determine the conventional (Maslen, Fox & O'Keefe, 2006) interpolating functions (4) and (5) for $0 \le \sin\theta/\lambda \le 2\ Å^{-1}$ and $2 \le \sin\theta/\lambda \le 6\ Å^{-1}$ ranges, respectively,

which will allow the users [should there be interest] include the expansions into X-ray diffraction software with only minor modifications;

c) optimize the extended interpolating functions (4) and (5) ($m = 5$ and $n = 4$, respectively) in order to increase the accuracy of the interpolated X-ray scattering factors.

1.7 Two-article dissertation structure

For this dissertation, an article-style format shall be used. It includes an introduction (Chapter1) which gives a general overview of the research interest and motivation and identifies gaps in the existing knowledge that our work trying to fill. Chapters 2 and 3 include the two research papers on the relativistic Dirac-Hartree-Fock X-ray scattering factors.

The first paper titled "*Revisited Relativistic Dirac-Hartree-Fock X-ray Scattering Factors For Ground State Neutral Atoms with Z = 2 – 118, using wavefunctions from B-spline Dirac-Hartree-Fock program"* (Olukayode, Fischer &Volkov, 2022) introduces our approach and includes a calculation of the fully relativistic Dirac-Hartree-Fock X-ray scattering factors for all neutral atoms with $Z = 2 – 118$. The paper has been submitted to *Acta Crystallographica Section A: Foundations and Advances* in July of 2022. In September of 2022, we have received the two referee reports that suggest publication of the paper after a "minor revision". That work is now in progress.

The second paper titled "*Revisited Relativistic Dirac-Hartree-Fock X-ray Scattering Factors. II. Chemically-Relevant Cations and Selected Monovalent Anions for Z = 2 – 112*" (Olukayode, Fischer &Volkov, 2023) has been written and shall be submitted immediately after the acceptance of the first paper. It is dedicated to the calculation of the fully relativistic Dirac-Hartree-Fock X-ray scattering factors for a total of 318 species that include a) all cations of the elements with $Z = 3$ (Li) – 104 (Rf) listed in Greenwood & Earnshaw (1997) (Figure 2.5,

page 28), b) six monovalent anions (O^{-}, F^{-}, Cl^{-}, Br^{-}, I^{-}, At^{-}), c) the ns^1np^3 excited (valence) states of carbon and silicon, and d) several exotic cations (Db^{5+}, Sg^{6+}, Bh^{7+}, Hs^{8+} and Cn^{2+}).

The final chapter (Chapter 4) offers a set of concluding remarks, a summary of the results obtained, and some of the future plans.

References

Atkins, P. & de Paula, J. (2006). *Atkins' Physical Chemistry*, 8th ed. W.H. Freeman and Company, New York.

Autschbach, J. (2012). *J. Chem. Phys*. **136**, 150902-1 – 150902-15.

Bunge, C. F., Barrientos, J. A. & Bunge, A. V. (1993). *At. Data Nucl. Data Tables*, **53**, 113 – 162.

Clementi, E. & Roetti, C. (1974). *At. Data Nucl. Data Tables*, 14, 177 – 478.

Clementi, E. & Raimondi, D. L. (1963). *IBM Res. Note NJ-27*.

Coppens, P. (1997). *X-ray Charge Density and Chemical Bonding*. Oxford University Press, New York.

Coulthard, M. A. (1967). *Proc. Phys. Soc.* **91**, 44 – 49.

Cromer, D. T. & Mann, J. B. (1968a). *X-ray scattering factors computed from numerical Hartree-Fock wave functions*. Los Alamos Scientific Laboratory Report LA-3816.

Cromer, D. T. & Mann, J. B. (1968b). *Acta Cryst*. A**24**, 321 – 324.

Cromer, D. T. & Waber, J. T. (1965). *Acta Cryst*. **18**, 104 – 109.

Cromer, D. T. & Waber, J. T. (1968). Unpublished work reported in *International Tables for X-ray Crystallography* (1974), Vol. IV, p. 71. Birmingham: Kynoch Press. (Present distributor: Kluwer Academic Publishers, Dordrecht.)

de Boor, C. (1977). *SIAM J. Numerical Analysis* **14**, 441 – 472.

Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* A**24**, 390 – 307.

Drake, G. W. F. & Goldman, S. P. (1981). *Phys. Rev. A* **23**, 2093 – 2098.

Dyall K. G., Grant, I. P., Johnson, C., Parpia, F.A. & Plummer, E. (1989). *Comput. Phys. Comm*. **55**, 425 – 456.

Feil, D. (1977). *Diffraction Physics. Isr. J. Chem* **17**, 103 – 110

Froese Fischer, C. (2007). *Adv. At. Mol. Opt. Phys*. **55**, 539.

Froese Fischer, C. (2011). *Comput. Phys. Comm.* **182**,1315 – 1326.

Froese Fischer, C. & Zatsarinny, O. (2009). *Comput. Phys. Comm.* **180**, 879 – 886.

Giacovazzo, C., Monaco, H. L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, G. & Catti, M. (1992). *Fundamentals of Crystallography*. Oxford University Press, New York.

Grant, I. P., McKenzle, B. J., Norrington, P. H., Mayers, D. F. & Pyper, N. C. (1980). *Comp. Phys. Comm.* **21**, 207 – 231.

Guerra, M., Amaroa, P., Santos, J. P. & Indelicato, P. (2017). *At. Nucl. Data. Tabl.* **117-118**, 439 – 457.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Kaldor, U. & Wilson, S. (2003). *Theoretical Chemistry and Physics of Heavy and Super heavy Elements, Progress in Theoretical Chemistry and Physics* Vol. 11

Koga, T., Kanayama, K., Watanabe, S. & Takkar, A. J. (1999). *Int. J. Quantum Chem.* **71**, 491 – 497.

Kołos & Wolniewicz (1964), *J. Chem. Phys.* **41**, 3663 – 3673.

Kołos & Wolniewicz (1965), *J. Chem. Phys.* **43**, 2429 – 2441.

Kołos & Wolniewicz (1968), *J. Chem. Phys.* **49**, 404 – 410.

Ladd, M. & Palmer, R. (2013). *Structure Determination by X-ray Crystallography*: *Analysis by X-rays and Neutrons*. 5th Edition, Springer.

Macchi, P. & Coppens, P. (2001). *Acta Cryst.* A**57**, 656 – 662.

Mann, J. B. (1968). Los Alamos Scientific Laboratory Report LA-3961.

Maslen, E. N., Fox, A. G. & O'Keefe, M. A. (1992). *International Tables for Crystallography, Vol. C*, edited by A. J. C. Wilson, pp. 476 – 511.

Michauk, C. & J. Gauss (2007). *J. Chem. Phys*. **127**, 044106.

Olukayode, S., Froese Fischer, C. & Volkov, A. (2022) *Acta Cryst.* A, being revised.

Olukayode, S., Froese Fischer, C. & Volkov, A. (2023) *Acta Cryst.* A, ready to be submitted.

Parpia, F. A., Froese, F. C. & Grant, I. P. (1996). *Comput. Phys. Commun.* **94**, 249 – 271.

Rez, D., Rez, P. & Grant, I. (1994). *Acta Cryst*. A**50**, 481 – 497.

Schwarzenbach, D. (1996). *Crystallography*. Wiley, Chichester.

Stout, G.H. & Jensen, L.H. (1989). *X-Ray Structure Determination: A Practical Guide,* Second Edition New York, NY.

Su, Z. & Coppens, P. (1997). *Acta Cryst.* A**53**, 749 – 762.

Vand, V., Eiland, P. F. & Pepinsky, R. (1957). *Acta Cryst*. **10**, 303 – 306

Wang, J., Smith, V. H., Bunge, C. & Jáiregui , R. (1996). *Acta Cryst*. A**52**, 649 – 658

Watson, R. E. (1958). *Phys. Rev.* **111**, 1108 – 1110.

Weissbluth, M. (1978). *Atoms and Molecules*. Academic Press, Inc., New York / London.

Zatsarinny, O.; Froese Fischer, C. (2016). *Comput. Phys. Comm.* **202**, 287 – 303. https://github.com/compas/dbsr_hf

Chapter 2. ARTICLE 1 - Revisited Relativistic Dirac-Hartree-Fock X-ray Scattering Factors. I. Neutral Atoms with Z = 2 – 118.

Authors

Shiroye Olukayode[a]*, Charlotte Froese Fischer[b] and Anatoliy Volkov[a]*

[a]Department of Chemistry and Computational Science Program, Middle Tennessee State University, Murfreesboro, TN, 37132, USA
[b]Department of Computer Science, University of British Columbia, 2366 Main Mall, Vancouver, British Columbia, V6T1Z4, Canada

Correspondence email: oos2b@mtmail.mtsu.edu; anatoliy.volkov@mtsu.edu

This paper is dedicated to the memory of Dr. Oleg Zatsarinny (1953-2021)

Synopsis Using relativistic wavefunctions from a B-spline Dirac-Hartree-Fock program DBSR_HF [Zatsarinny & Froese Fischer (2016). *Comput. Phys. Comm.* **202**, 287 – 303] we have re-determined the X-ray scattering factors and the corresponding interpolating approximations in the 0 – 2 and 2 – 6 Å^{-1} $\sin\theta/\lambda$ ranges for all neutral atoms between He (Z=2) and Og (Z=118). The results seem to represent an excellent compromise among all the previous studies and should (hopefully) be free of typos and inconsistencies. The generated data have been tabulated using the same format and expansions (as well as more accurate extended expansions) as used in the 2006 edition of Volume C of *International Tables for Crystallography*, and thus can be readily incorporated into the existing X-ray diffraction software

Abstract In this first of a series of publications, we revisit the X-ray scattering factors for neutral atoms. Using the recently developed DBSR_HF program [Zatsarinny & Froese Fischer (2016). *Comput. Phys. Comm.* **202**, 287 – 303] we have calculated the fully relativistic Dirac-Hartree-Fock ground-state wavefunctions for all atoms with Z = 2 – 118 (He – Og) using the extended average level scheme and including both the Breit interaction correction to the electronic motion due to magnetic and retardation effects, and the Fermi distribution function for the description of the nuclear charge density. The comparison of our wavefunctions with those obtained in several previous studies in terms of the total and orbital (spinor) energies, and a number of local and integrated total and orbital properties, confirmed the quality of the generated wavefunctions. The employed dense radial grid combined with the DBSR_HF's B-spline representation of the relativistic one-electron orbitals allowed for a precise integration of the X-ray scattering factors using a newly developed Fortran program SF. Following the established procedure [Maslen, Fox & O'Keefe (2006) in *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589], the resulting X-ray scattering factors have been interpolated in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} ranges using the recommended analytical functions with both the four- (which is a current convention) and five-term expansions. An exhaustive comparison of the newly generated X-ray scattering factors with the IUCr-recommended values and those from a number of previous studies showed an overall good agreement and allowed us to identify a number of typos and inconsistencies in the recommended quantities. A detailed analysis of the results suggests that the newly derived values may represent an excellent compromise among all the previous studies. The determined conventional interpolating functions for the two $\sin\theta/\lambda$ intervals show, on average, the same accuracy as the recommended parametrizations. However, an extension of each expansion by only a single term provides a significant improvement in the accuracy of the interpolated values for an overwhelming majority of the atoms. As such an updated set of the fully relativistic X-

ray scattering factors and the interpolating functions for neutral atoms with $Z = 2 – 118$ can be easily incorporated into the existing X-ray diffraction software with only minor modifications. The outcomes of the undertaken research should be of interest to the members of crystallographic community who push the boundaries of the accuracy and precision of the X-ray diffraction studies.

Keywords: **Relativistic X-ray scattering factors for neutral atoms; relativistic Dirac-Hartree-Fock; interpolation of X-ray scattering factors.**

2.1 Introduction

Since the pioneering work of Sir William Henry Bragg (1862-1942) and his son Sir William Lawrence Bragg (1890-1971) that was recognized by their joint 1915 Nobel Prize in Physics, the X-ray crystal diffraction has become one of the most accurate, precise, and perhaps the most commonly used technique for studying the crystalline state of matter at the atomic scale. At the heart of the method is the atomic X-ray scattering (form) factor (XRSF) defined as a Fourier transform of the atomic electron density $\rho(\mathbf{r})$ (Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997):

$$f(\mathbf{S}) = \int \rho(\mathbf{r}) e^{2\pi \mathrm{i}\mathbf{S}\cdot\mathbf{r}} \mathrm{d}^3\mathbf{r} \tag{28}$$

where $\mathbf{r}$ is a vector in direct space given as $\mathbf{r} \equiv (r, \theta, \phi)$ in the spherical coordinate system, $\mathbf{S}$ is the so-called scattering vector, i is the imaginary unit ($\mathrm{i} = \sqrt{-1}$) and $\rho(\mathbf{r}) = \rho(r, \theta, \phi)$ is the electron density of an atom (or a particular electronic shell). The scattering vector $\mathbf{S}$ has the length $S = |\mathbf{S}| = 2 \sin\theta / \lambda$, where the angle θ is defined in such a way that 2θ is the angle between the directions of the incident and diffracted X-rays (Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997), and λ is the wavelength of the incident X-ray beam. Following Giacovazzo et al. (1992) and Coppens (1997), we assume that $\mathbf{S}$ is directed along the z-axis of the spherical coordinate system, in which case

$$\mathbf{S} \cdot \mathbf{r} = S\, r \cos\theta = 2 \sin(\theta) / \lambda\, r \cos\theta = 2\, s\, r \cos\theta$$

where $s = \sin\theta / \lambda = S/2$. Then, the integral (28) can be re-written as

$$f(s) = \int_0^{\infty} \int_0^{2\pi} \int_0^{\pi} \rho(r, \theta, \phi) e^{4\pi \mathrm{i} s r \cos\theta} r^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi \, \mathrm{d}r \tag{29}$$

Using the non-relativistic quantum mechanical treatment of an atom, the electron density can be expressed in terms of the atomic orbital wavefunctions $\psi(r, \theta, \phi)$

$$\rho(r,\theta,\phi) = \sum_{i=1}^{N_{\mathrm{occ}}} n_i |\psi_i(r,\theta,\phi)|^2 \tag{30}$$

where ψ_i and n_i are the wavefunction and the electron population of the *i*-th orbital, respectively, N_{occ} is the number of the occupied orbitals, and the square modulus $|x|^2$ is equal to x^2 if the quantity x is real and x^*x if it is complex (symbol * denotes the complex-conjugate). Representing wavefunction of the *i*-th atomic orbital, $\psi_i(r,\theta,\phi)$, as a product of the radial function $R_i(r)$, which is always real and normalized,

$$\int_0^\infty R_i^2(r) r^2 \,\mathrm{d}r = 1 \tag{31}$$

, and the complex spherical harmonic angular function $Y_i(\theta,\phi)$ (Weissbluth, 1978), the expression for the electron density becomes

$$\rho(r,\theta,\phi) = \sum_{i=1}^{N_{\mathrm{occ}}} n_i R_i^2(r) |Y_i(\theta,\phi)|^2 \tag{32}$$

For a spherically symmetric atom, which appears to be an excellent first-order approximation (Giacovazzo et al., 1992; Schwarzenbach, 1996; Coppens, 1997), the angular part is just

$$Y(\theta,\phi) = (4\pi)^{-1/2} \tag{33}$$

Introducing a simplified notation for the radial part of the non-relativistic atomic electron density

$$\mathbb{R}(r) = r^2 \sum_{i=1}^{N_{\mathrm{occ}}} n_i R_i^2(r) \tag{34}$$

, inserting the definitions (34), (33), and (32) into the integral (29), and integrating over θ and ϕ, the final expression for the atomic XRSF becomes:

$$f(s) = \frac{1}{4\pi} \int_0^\infty \mathbb{R}(r) \frac{\sin(4\pi s r)}{sr} dr \tag{35}$$

In relativistic quantum mechanics (Swirles, 1935; Grant, 1961, 1970) the radial part of the relativistic *i*-th orbital (spinor) includes two components: the large component $P_i(r)$ and a small component $Q_i(r)$ normalized as (Grant, 1961)

$$\int_0^\infty [P_i(r)^2 + Q_i(r)^2\,]\,\mathrm{d}r = 1 \tag{36}$$

As such, the atomic XRSF based on relativistic wavefunction for a spherically symmetric atom is

$$f(s) = \frac{1}{4\pi}\int_0^\infty \mathcal{R}(r)\frac{\sin(4\pi sr)}{sr}\,dr \tag{37}$$

where

$$\mathcal{R}(r) = \sum_{i=1}^{N_{\mathrm{occ}}} n_i[P_i(r)^2 + Q_i(r)^2\,] \tag{38}$$

The X-ray scattering factors for neutral atoms with Z = 2 (He) – 98 (Cf) recommended by IUCr (Maslen, Fox & O'Keefe, 2006) were calculated at the single-configuration relativistic Hartree-Fock (Dirac-Hartree-Fock, DHF) level (Swirles, 1935; Grant, 1961, 1970) by Doyle & Turner (1968) based on wavefunctions of Coulthard (1967), and by Cromer & Waber (1968) using the wavefunctions of Mann (1968). The Coulthard (1967) approach included the point-charge model while Mann (1968) accounted for a finite charge distribution of the atomic nucleus, and both neglected the magnetic and retardation effects. The original Cromer & Waber (1968) XRSFs that terminated at $\sin\theta/\lambda = 2$ Å^{-1} were extrapolated to 6 Å^{-1} by Fox, O'Keefe & Tabbernor (1989). The Doyle & Turner (1968) scattering factors are labelled as RHF in Volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006), while the Cromer & Waber (1968) / Fox, O'Keefe & Tabbernor (1989) data are denoted as *RHF. The X-ray scattering for a neutral hydrogen atom were calculated from an exact non-relativistic quantum mechanical treatment (Maslen, Fox & O'Keefe, 2006).

The conventional interpolations of XRSFs (Maslen, Fox & O'Keefe, 2006) include the function (Vand, Eiland & Pepinsky, 1957)

$$f(s) = \sum_{i=1}^{m} a_i \exp(-b_i s^2) + c \tag{39}$$

with $m = 4$ for interpolation of the scattering factors in the $0 \leq s \leq 2$ Å^{-1} interval (Doyle & Turner, 1968), and

$$f(s) = \exp\left(\sum_{i=0}^{n} a_i s^i\right) \tag{40}$$

with $n = 3$ for the $2 \leq s \leq 6$ Å^{-1} range (Fox, O'Keefe & Tabbernor, 1989).

The recommended XRSFs for chemically significant *ions* (Maslen, Fox & O'Keefe, 2006) were evaluated at different levels of theory, including the non-relativistic Hartree-Fock and correlated methods, and relativistic Dirac-Hartree-Fock and Dirac-Slater techniques.

The importance of including in quantum chemical calculations the relativistic effects that arise due to velocities of moving electrons approaching the speed of light have been emphasized, for example, in a series of recent reviews by Pyykkö (2012), Autschbach (2012), and Pyper (2020). A significant part of each one of these publications is dedicated to a discussion of the substantial differences in the electron density distribution between the relativistic and non-relativistic predictions, with relativistic ones pretty much always being closer or perfectly matching the experimental observations.

While there have been several attempts to improve XRSFs, in the following only the studies that include the relativistic calculations are briefly discussed.

In 1994, Rez, Rez & Grant (1994, 1997) redetermined XRSFs for naturally occurring elements and ions using multiconfiguration Dirac-Hartree-Fock combined with the extended average level (EAL) model (Grant, Mayers & Pyper, 1976) as was implemented in the Oxford MCP/MCDF package (Grant et al., 1980) and later included in the GRASP suite of programs

(Dyall et al., 1989). The authors found "*very little difference (less than 0.1%) between* [their] *results and those of Doyle & Turner (1968) except for some heavy elements at high values of* $sin\,\theta\,/\lambda$ " despite the expected differences between their multiconfiguration DHF wavefunctions and single-configuration DHF wavefunctions used by Doyle & Turner (1968) and Cromer & Waber (1968). For interpolation, Rez, Rez & Grant (1994, 1997) use parametrization (39) with $m = 4$ but excluded the constant c. The two interpolating functions were designed to provide a high accuracy fit for $0 \leq s \leq 2$ Å^{-1} and a lower accuracy fit for the entire 0 – 6 Å^{-1} range (Rez, Rez & Grant, 1994, 1997).

The 1996 calculations by Wang et al. (1996) were also performed at the multiconfiguration DHF level but were limited to neutral atoms with Z = 2 (He) – 18 (Ar). An important feature of the study was that the relativistic wavefunctions approached the ground-state non-relativistic wavefunctions as the speed of light approached the infinity. The authors also chose to generate scattering factors on a fine grid instead of providing analytical interpolating functions.

Su & Coppens (1997) argued that the Rez, Rez & Grant (1994, 1997) wavefunctions may not "*correspond to the optimized ground states of the atoms*" and offered their own XRSFs (Su & Coppens, 1997, 1998) for neutral atoms with Z = 1 – 54 (Xe) computed using the multiconfiguration DHF method with the optimal level (OL) model in GRASP92 (Parpia, Froese Fischer & Grant, 1996). In 2001, Macchi & Coppens (2001) extended the work of Su & Coppens (1997, 1998) to all chemically relevant ions up to I^-. It is a little surprising that Volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) does not include the values from those two studies. Perhaps, it is related to the fact that Su & Coppens (1997, 1998) and Macchi & Coppens (2001) did not process all the atoms and ions listed in the currently used tables (Maslen, Fox & O'Keefe, 2006), or because the interpolating procedure for XRSFs employed in their studies was somewhat different from

the established approach: they used the function (39) with $m = 6$ and $c = 0$ in each of the 0 – 2, 2 – 4 and 4 – 6 $Å^{-1}$ $\sin\theta/\lambda$ intervals. Even after the passing of Prof. Coppens in 2017, the scattering factors from these studies are available online at http://harker.chem.buffalo.edu/group/ptable.html. However, one should be aware that for several atoms some parts of the data are missing.

In the last two decades, there has been a significant increase in the level and number of available software for relativistic atomic calculations and an enormous progress in computing power, to the point that accurate relativistic atomic calculations have become almost routine and can be performed on an average desktop or laptop computer. In addition to the atomic relativistic codes with a long history that are still being supported and improved, such as for example, MCDFGME (Desclaux, Mayers & O'Brien, 1971; Desclaux, 1975; Indelicato & Desclaux, 1990; Guerra et al., 2017) and GRASP (Grant et al., 1980; Dyall et al., 1989; Jönsson et al., 2007; Jönsson et al., 2013; Froese Fischer et al., 2019), there have been a number of new codes developed including DFRATOM (Matsuoka & Watanabe, 2001), FAC (Gu, 2008), AMBiT (Kahl & Berengut, 2019), and DBSR_HF (Zatsarinny & Froese Fischer, 2016).

A general Dirac-Hartree-Fock program DBSR_HF written by Zatsarinny & Froese Fischer (2016) uses a B-spline basis (Bachau et al., 2001; Froese Fischer, 2007, 2021) for the large and small components of the relativistic orbitals (spinors) which makes it ideal for studies involving the electron density distribution as the electron density at any point in space can be calculated very precisely directly from a B-spline representation without additional interpolation, and when combined with a fine grid, allows a straightforward evaluation with a high precision of a pretty much any integral that involves the electron density. The Breit interaction model which accounts for the relativistic corrections to the electronic motion due to magnetic and retardation effects can be included in the setup of the Hamiltonian matrix (Zatsarinny & Froese Fischer, 2016) which makes the effect present in the calculated

wavefunctions. The program supports several nuclear models including the point charge-based Coulomb function, and two schemes that account for the finite size of an atomic nucleus: the uniform ball of charge model, and the Fermi distribution function (Johnson & Soff, 1985; Zatsarinny & Froese Fischer, 2016). The built-in utility allows for a flexible energy functional input: the code can automatically expand a specified non-relativistic configuration into a list of corresponding relativistic configuration state functions (CSFs) with appropriate statistical weights. Alternatively, a user can create his/her own list of relativistic CSFs with custom weights. The DBSR_HF program has a simple, keyword-driven input and a user-friendly output. The former is especially useful when calculating the atomic excited states. The code is written in a plain Fortran90 and is well commented which makes it easy to modify when a custom output or interface is needed. Perhaps, the DBSR_HF's only weakness is that while the quantum electrodynamical (QED) corrections due the self-energy of the electrons and vacuum polarization, evaluated using the same approximations as in GRASP (Jönsson et al, 2007), can be included in the calculations of energies, due to theoretical limitations of the methods implemented in DSBR_HF, they cannot be applied to the resulting one-electron wavefunctions (Zatsarinny & Froese Fischer, 2016). Dr. Zatsarinny passed away before he was able to extend the functionality of the code.

In this first of a series of papers, we i) calculate the Dirac-Hartree-Fock ground-state wavefunctions of neutral atoms with Z = 2 (He) – 118 (Og) using the DBSR_HF code of Zatsarinny & Froese Fischer (2016), ii) evaluate the X-ray scattering factors from the resulting one-electron densities represented in a B-spline basis, and iii) determine the coefficients of the interpolating functions following the same approach as given in *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). For those users who require higher accuracy interpolation, we provide the extended expansions within the same general formalism. In a future study, we plan to apply the described approach to all chemically relevant ions so that the

X-ray scattering factors for all neutral atoms with $Z = 2 – 118$ and their ions will be determined at the relativistic Dirac-Hartree-Fock level of theory in a consistent manner.

2.2 Methods

2.2.1 Relativistic calculations

The relativistic atomic calculations at the Dirac-Hartree-Fock level were performed using the DBSR_HF program (Zatsarinny & Froese Fischer, 2016) that was downloaded from the GitHub repository (https://github.com/compas/dbsr_hf). The code was slightly modified to a) process atoms up to $Z = 118$ (Og), b) include up-to-date physical constants and weights of the most abundant or most stable isotopes (CRC Handbook of Chemistry and Physics, 2021), and c) provide an interface to the newly written Fortran code SF (see description below) that calculates XRSFs, determines coefficients of the interpolating functions, and calculates a number of integrated local and atomic properties. Finally, a newly written Fortran program ANALYSIS collects the SF output files, performs statistical analysis, compares results with those from the previous studies, and prints out a number of tables in a user-friendly format. A flow chart that shows the sequence of steps and the associated software used is shown in Figure 1.

The ground-state electron configurations for all atoms were specified in the non-relativistic LS notation (Table S1). For atoms with $Z = 2$ (He) – 104 (Rf) the electron configurations were taken from the latest edition of *CRC Handbook of Chemistry and Physics* (Martin, 2021), while for heavier elements we used the proposed configurations based on those in the 6th period of the Mendeleev Periodic Table (WebElements, 2022). The ground-state electron configurations employed in our work are in complete agreement with those used in a recent study by Guerra et al. (2017). In contrast, Tatewaki, Yamamoto & Hatano (2017), who limited their calculations to $Z = 1 – 103$ (Lr), used different electron configurations for Bk and Lr, as shown in Table S1 and summarized below:

Bk (Z=97)	[Rn] $5f^9\ 7s^2$	Visscher & Dyall (1997), Guerra et al. (2017), this work
	[Rn] $5f^8\ 6d\ 7s^2$	Tatewaki et al. (2017)
Lr (Z=103)	[Rn] $5f^{14}\ 7s^2\ 7p$	Guerra et al. (2017), this work
	[Rn] $5f^{14}\ 6d\ 7s^2$	Visscher & Dyall (1997), Tatewaki et al. (2017)

Visscher & Dyall (1997) whose calculations extend to Mt (Z=109) also used the [Rn] $5f^{14}\ 6d\ 7s^2$ configuration for Lr. However, the electronic ground state of Lr was determined to be [Rn] $5f^{14}\ 7s^2\ 7p$ by Desclaux & Fricke (1980) using relativistic calculations though Martin (2021) marks it (together with [Rn] $5f^{14}\ 6d^2\ 7s^2$ for Rf) as "uncertain".

For a given non-relativistic LS configuration, the DBSR_HF code automatically generates all the appropriate relativistic configuration state function (CSFs) with the corresponding statistical weights which are listed in Table S2. The statistical weights are defined with respect to the number of Slater determinants used in the configuration expansion (Desclaux, 1973; Zatsarinny & Froese Fischer, 2016). For example, the ground state electronic configuration of sulfur (Z=16) $[Ne]3s^2 3p^4$ is expanded in terms of three CSFs:

$$3s^2\ 3p_{1/2}^2\ 3p_{3/2}^2 \qquad w = 0.40000$$

$$3s^2\ 3p_{1/2}^1\ 3p_{3/2}^3 \qquad w = 0.53333$$

$$3s^2\ 3p_{3/2}^4 \qquad w = 0.06667$$

where the subscript of the orbital symbol denotes the quantum number j ($j = l \pm s$) and w is the CSF weight. For, chlorine (Z=17, $[Ne]3s^2 3p^5$), there are two CSFs:

$$3s^2\ 3p_{1/2}^2\ 3p_{3/2}^3 \qquad w = 0.66667$$

$$3s^2\ 3p_{1/2}^1\ 3p_{3/2}^4 \qquad w = 0.33333$$

while for argon (Z=18), a closed-shell atom with the ground state electronic configuration $[Ne]3s^2 3p^6$, there is only one CSF:

$$3s^2\ 3p_{1/2}^2\ 3p_{3/2}^4 \qquad w = 1.00000$$

In that respect, our calculations are similar to the average level (AL) and extended average level (EAL) formalisms developed by Grant, Mayers & Pyper (1976) and incorporated in GRASP (Grant et al., 1980; Dyall et al., 1989). Even though Su & Coppens (1997) argued that "*multiconfiguration wavefunctions from the average level model […] do not correspond to the optimized ground states of the atoms*", the EAL formalism was used by Rez, Rez and Grant (1994) in their work on XRSFs. The AL approach was also employed by Visscher & Dyall (1997) who performed calculations for the first 109 elements in GRASP thus providing a benchmark set of the total electronic ground-state energies, as well as the radial expectation values and energies of the spinors. Unfortunately, the spinor parameters were not included in the paper and instead were made available at "http://theochem.chem.rug.nl/~luuk/FiniteNuclei/" that is no longer accessible. As stated in Dyall et al. (1989), "[t]*he object of the E/AL options is to determine a set of orbitals which are optimal for the average energy of a set of CSFs*", which from our point of view, provides a more realistic description for the electron distribution of an atom in a crystal as due to bonding effects, it will not be always confined to the lowest energy level of the ground state configuration that is optimized in the OL method: "[a]*n OL calculation produces only one reliable energy level, namely the one being optimized*" (Grant et al., 1980).

The DBSR_HF calculations were performed using the default values for the order of B-splines (which is equal to 9), energy convergence tolerance of 1.0×10^{-10} atomic units (a.u.), and the orbital convergence tolerance of 1.0×10^{-7} a.u. In contrast to the original code (Zatsarinny & Froese Fischer, 2016) in which the maximum number of iterations was set at 25, in the latest version of DBSR_HF that number is increased to 75. In order to achieve a desired precision of the resulting wavefunctions (see discussion below) we significantly increased the density of the default "semi-logarithmic" grid (Zatsarinny & Froese Fischer, 2016) relative to the default settings. It involved a reduction of the initial step h_i (specified in atomic units) and

the factor h_e that is responsible for exponential growth of the grid in the "middle-radii region" from their default values of 0.25 to 0.05, and the reduction of the maximum step size at "high radii" from 1 to 0.5 atomic units.

The calculated wavefunctions include the Breit interaction thus accounting for the relativistic corrections to the electronic motion due to magnetic and retardation effects ("mbreit=2" option in DBSR_HF). We note that the original code developed by Coulthard (1967) which was used to generate the wavefunctions from which Doyle & Turner (1968) calculated the IUCr-recommended XRSFs (Maslen, Fox & O'Keefe, 2006) neglected not only the magnetic interactions between electrons (which can be simulated in DBSR_HF using the "mbreit=0" option) but also the finite radius of the nucleus (Coulthard, 1967). There are very few details about the relativistic calculations of Mann (1968), whose wavefunctions were used by Cromer & Waber (1968), except that the approach included "*the potential due to a finite nucleus, rather than a point nucleus*". It is likely that Mann's (1968) calculations included the homogeneously charged sphere model of a nucleus (Desclaux, 1973). In contrast, our DBSR_HF calculations employed the Fermi distribution function ("nuclear=Fermi" option) for the description of the nuclear charge density (Johnson & Soff, 1985; Zatsarinny & Froese Fischer, 2016). However, as mentioned in the Introduction section, the quantum electrodynamical (QED) corrections due the self-energy of the electrons and vacuum polarization were included in the energies only.

The convergence issues encountered for Cu, Pd, Tm and Yb were solved by first completing the calculation in a different state (Cu – [Ar] $3d^9\ 4s^2$, Pd – [Kr] $4d^9\ 5s^1$, Tm – [Xe] $4f^{12}\ 5d^1\ 6s^2$, Yb – [Xe] $4f^{13}\ 5d^1\ 6s^2$) and then using wavefunctions from those calculations as a starting point for the ground-state calculations: Cu – [Ar] $3d^{10}\ 4s^1$, Pd – [Kr] $4d^{10}$, Tm – [Xe] $4f^{13}\ 6s^2$, Yb – [Xe] $4f^{14}\ 6s^2$ (Table S1). In addition, the orbital convergence tolerance for the

ground-state calculations of Cu and Yb had to be increased to 5.0×10^{-7} and 1.1×10^{-7} a.u., respectively.

2.2.2 X-ray scattering factors calculations

The X-ray scattering factors were calculated in a newly developed Fortran code SF (Figure 1) that is interfaced to DBSR_HF via a binary file that contains all the information necessary to work with relativistic one-electron orbitals in the B-spline representation. For consistency, the SF code includes DBSR_HF's subroutines "BVALU2" and "INTERV1", which are the modifications of de Boor's subroutines "BVALU" and "INTRV" (de Boor, 1977), that are used to evaluate a B-spline and its derivatives at any point.

The integration of XRSFs and other properties (such as the mean orbital and atomic radii) in SF can be performed on a fixed Gaussian-type quadrature which employs the same abscissas and weights as used in DBSR_HF, or with the help of subroutines GAUS8 or QAG, both taken from the SLATEC package (Vandevender & Haskell, 1982; Fong et al., 1993). GAUS8 (Jones, 1981) is an automatic numerical integrator of a single variable that uses an adaptive 8-point Legendre-Gauss algorithm and is designed "*for high accuracy integration or integration of smooth functions*" (Jones, 1981). The subroutine QAG (Piessens & de Doncker, 1980) serves the same purpose and shares a similar philosophy with GAUS8, but is globally adaptive and for a local integration it uses the Gauss-Kronrod rule with a user-defined number of points (7 – 61). The exhaustive numerical tests showed that a fine radial grid had to be employed in DBSR_HF in order to achieve the desired agreement of at least eight decimal digits in XRSFs for heavy atoms at high $\sin\theta/\lambda$ values (at or above $3-4$ Å^{-1}) between the two adaptive integrators and DBSR_H's fixed-points quadrature. This was the main reason for significantly increasing the density of the radial grid in the DBSR_HF calculations relative to the default settings. In this work, all integrations including orbital and atomic mean radii, and the XRSFs calculations were performed using subroutine QAG (Piessens & de Doncker, 1980)

which also returns what appears to be a reliable estimate of the precision of the evaluated integral, though essentially the same results are obtained with the other two integrators.

The X-ray scattering factors for each atom were generated using the $\sin\theta/\lambda$ grid proposed by Wang et al. (1996) and used in Table 6.1.1.1 of Volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). The fitting (described in the following section) in the $0 \leq s \leq 2$ Å^{-1} range included the same $\sin\theta/\lambda$ grid as used in Doyle & Turner (1968). For fitting in the $2 \leq s \leq 6$ Å^{-1} range our approach was different from the one described by Fox, O'Keefe & Tabbernor (1989): we generated 41 × scattering factors with $s = \sin\theta/\lambda = 2.0,\ 2.1,\ 2.2 \ldots 6.0$ Å^{-1} and used those to optimize the coefficients in equation (40). However, the user of the SF code has a full control over the $\sin\theta/\lambda$ grid parameters which can be conveniently specified in an input file.

2.2.3 X-ray scattering factor interpolating functions

Interpolation of the integrated XRSFs employed the same expansions as given in section 6.1 of volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). For $0 \leq s \leq 2$ Å^{-1} we used the function (39) (Vand, Eiland & Pepinsky, 1957) in which the parameter m was set to 4 as in Doyle & Turner (1968) and Maslen, Fox & O'Keefe (2006), and then extended to $m = 5$. In the expansion (40) used to fit the scattering factors in $2 \leq s \leq 6$ Å^{-1}, the fits were performed with $n = 3$ following Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006), and then extended to $n = 4$. We find these choices to be more practical than the six-Gaussian expansions in three different $\sin\theta/\lambda$ regions used in Su & Coppens (1997, 1998) and Macchi & Coppens (2001) since the results can be incorporated in the currently used X-ray structure refinement software with very minor modifications. That said, the fitting schemes used in our code are very general and can handle any number of terms in each of the expansions (39) and (40).

In the original work by Doyle & Turner (1968) and a later study by Rez, Rez & Grant (1994, 1997) the fitting was done using the Marquardt method (Marquardt, 1963; Bevington & Robinson, 2002) which also known as the Levenberg-Marquardt method (Levenberg, 1944; Fletcher, 1987; Press et al., 2007). In brief, the Levenberg-Marquardt algorithm adaptively varies the parameter updates between the gradient (steepest) descent update and the Gauss-Newton update. Fox, O'Keefe & Tabbernor (1989) for their high $\sin\theta/\lambda$ extrapolations and interpolations used a proprietary Apple Macintosh software "Cricket Graph" (Rafferty & Norling, 1986). The studies by Su & Coppens (1997, 1998) and Macchi & Coppens (2001) employed the L-BFGS-B optimization code (Zhu et al. 1994; Byrd et al., 1995) which also enables one to constrain the fitted parameter within the user-specified bounds.

The non-linear least squares fitting protocol incorporated in our Fortran code SF is described in detail in Appendix A.

2.3 Assessment of the quality of the calculated wavefunctions

The quality of the performed DBSR_HF calculations was a subject of a rigorous comparison in terms of the i) total electronic energies, ii) atomic radii, and iii) multiple properties of the orbitals (spinors) such as the electronic energy, mean radius, and location of the charge density maximum, with a number of previous studies (Wang et al., 1996; Su & Coppens, 1997; Visscher & Dyall, 1997; Tatewaki, Yamamoto & Hatano, 2017; Guerra et al., 2017). The outcome of this analysis, presented in Appendix B, confirms the quality of the generated wavefunctions.

2.4 X-ray scattering factors and interpolations

The discussion of the newly generated X-ray scattering factors is separated into three sections. In section 4.1 we describe the differences between the DBSR_HF-based XRSFs for neutral atoms within the $0 \leq s \leq 6$ Å^{-1} range calculated in this study relative to the ones included in

the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) that are based on the works of

a) Doyle & Turner (1968) who used the wavefunctions of Coulthard (1967); these are designated as RHF in Table 6.1.1.1 in Maslen, Fox & O'Keefe (2006),
b) Cromer & Waber (1968) who used the wavefunctions of Mann (1968), with the extrapolations to 6 Å^{-1} performed by Fox, O'Keefe & Tabbernor (1989); these are labelled as *RHF in Table 6.1.1.1 in Maslen, Fox & O'Keefe (2006).

, as well as with those from the later studies by Rez, Rez & Grant (1994), Wang et al. (1996), Su & Coppens (1997) / Macchi & Coppens (2001). We note that the acronym "RHF" used in *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) is somewhat misleading as in quantum chemistry it usually refers to a "restricted Hartree-Fock" calculation. We think that a "DHF" (Dirac-Hartree-Fock) abbreviation is more appropriate.

In sections 4.1 and 4.2 we present the results of the analytical interpolations of the new XRSFs in the $0 \leq s \leq 2$ Å^{-1} and $2 \leq s \leq 6$ Å^{-1} intervals, respectively. As mentioned above, equation (39) with $m = 4$ and $m = 5$ was used to fit the scattering factors in the $0 \leq s \leq 2$ Å^{-1} range, while the scattering factors in the $2 \leq s \leq 6$ Å^{-1} interval were fitted via equation (40) with $n = 3$ and $n = 4$.

2.4.1 X-ray scattering factors

The X-ray scattering factors for neutral atoms with Z = 2 (He) – 118 (Og) are listed in Table S3 using the Wang et al. (1996) $\sin\theta/\lambda$ grid and the same format as in Maslen, Fox & O'Keefe (2006). A detailed comparison of the two sets of values for the elements with Z = 2 – 98 (He – Cf) is given in Figures 2 – 11 in the main body of the manuscript and Figure S1 in the Supporting Information section using the $\sin\theta/\lambda$ grid of Doyle & Turner (1968). We note that while in Table S3 we report the new scattering factors to five decimal places, the precision of the calculated values should be at least eight decimal digits.

In Figures 2 – 11 and S1, and the discussion below we make use of the *absolute* differences,

$$\Delta f = |f_{\text{this work}} - f_{\text{benchmark}}| \tag{41}$$

, and the *relative* differences,

$$\Delta f_{\%} = \frac{|f_{\text{this work}} - f_{\text{benchmark}}|}{f_{\text{benchmark}}} \times 100\% \tag{42}$$

, between the two sets of scattering factors. In these equations, $f_{\text{benchmark}}$ represents the scattering factors from one of the benchmark studies that are labelled in the figures as follows: DT - Doyle & Turner (1968), CW - Cromer & Waber (1968), RRG - Rez, Rez & Grant (1994), WSBJ - Wang et al. (1996), SCM - Su & Coppens (1997) and Macchi & Coppens (2001). In several figures, the abbreviation DT/CW is used to denote the entire set of the Doyle & Turner (1968) and Cromer & Waber (1968) values as given in Maslen, Fox & O'Keefe (2006). In Figures 2 – 5 the maximum and mean Δf and $\Delta f_{\%}$ quantities are plotted as a function of the atomic number *Z*, and the mean differences were calculated for each atom over all the $\sin\theta/\lambda$ points in the Doyle & Turner (1968) grid. The graphs in Figures 6 – 11 show the same quantities plotted as a function of $\sin\theta/\lambda$ using the Doyle & Turner (1968) grid, and the averaging for each $\sin\theta/\lambda$ grid point was done over all specified atoms. Because it is cumbersome to plot all atoms on the same graph, each figure includes only one group of atoms. As such, it is convenient to discuss each group separately.

2.4.1.1 He (Z=2) –Ar (Z=18)

In Figures 2a, 3a, 4a and 5a we plot the maximum and mean differences Δf and $\Delta f_{\%}$ for the elements up to Ar (*Z*=18). Note that while in Figures 2a and 3a that show the maximum and mean Δf as a function of *Z* we start with He (*Z*=2), in Figures 4a and 5a that display the graphs of the maximum and mean $\Delta f_{\%}$, the first atom plotted is boron (*Z*=5). This is done in order to avoid unreasonably large percentile differences $\Delta f_{\%}$ for He, Li and Be that originate from a limited precision with which XRSFs are reported in some of the benchmark studies: DT give

the values to three decimal places, while RRG and WSBJ use four decimal digits. As evident from Figures 2a and 3a, for most of these atoms the agreement between our work and the benchmark data, especially with the DT and RRG studies, is very good. The largest maximum and mean Δf values are observed for atoms in the middle of the periods: C, N, O, Si, P, and S. The WSBJ and SCM values are very close to each other but show noticeable deviations for the atoms in the middle of the two periods from the DT, RRG and our values. These differences are likely related to a different approximation level used in the relativistic calculations: OL in WSBJ and SCM, and AL/EAL in DT, RRG and this study. When the differences Δf are plotted as a function of $\sin\theta/\lambda$ for each of these atoms (Figures S1-6 – S1-8 and S1-14 – S1-16) the differences are restricted to $\sin\theta/\lambda \leq 0.4$ Å^{-1} which is an indication that it is the valence charge density distribution effect. It is interesting to note that while for atoms in the middle of a period our results agree almost perfectly with the values reported in the RRG study yet deviate noticeably from the WSBJ and SCM data, at the beginning and end of each period the situation is opposite: our values are in a complete agreement with the WSBJ and SCM data, while the RRG numbers show noticeable deviations. The Δf values for the DT study pretty much always lie in between the RRG and SCM deviations.

In terms of the percentile XRSF differences ($\Delta f_{\%}$) (Figures 4a) the largest deviations of 1.5% and 2% are observed for the DT scattering factors at 6 Å^{-1} for the C and N atoms, respectively, and are due to a limited precision (three decimals) of the DT data. Note that He, Li and Be were excluded from these graphs because of that very same issue. However, Figures S1-1 – S1-3 show that for He, Li and Be all studies give essentially identical results. The mean values of $|\Delta f_{\%}|$ obtained by averaging over all $\sin\theta/\lambda$ grid points (Figure 5a), are below 0.5% for all studies and all atoms between B (Z=5) and Ar (Z=18). It is interesting to note that the mean $\Delta f_{\%}$ between our work and the DT study decreases steadily as the atomic number Z

increases. It makes sense: as Z increases the magnitude of XRSF increases as well, so the limited precision of the DT data becomes less relevant.

In Figures 6a, 7a, 8a and 9a we plot the maximum and mean Δf and $\Delta f_{\%}$ values as a function of $\sin\theta/\lambda$ where the averaging at each $\sin\theta/\lambda$ grid point was done over all atoms between B (Z=5) and Ar (Z=18). In Figures 6a and 8a we also list atoms for which the maximum Δf and $\Delta f_{\%}$ values were observed at each $\sin\theta/\lambda$ grid point. In agreement with the trends seen in Figures S1-2 – S1-18, the WSBJ and SCM values are very close to each other but for several atoms deviate noticeably from the other studies at $\sin\theta/\lambda \leq 0.4$ Å^{-1}. This is likely due to differences in the valence density distributions between the OL and EAL wavefunctions. However, on a relative scale, these differences are very small: below 1% for the maximum and 0.2% for the mean $\Delta f_{\%}$ values. The RRG data agree with our values very well, while DT study shows slightly larger overall deviations. The unusually large mean and maximum percentile differences in the DT study at high $\sin\theta/\lambda$ (Figures 8a and 9a) are due to a limited precision (three decimal digits) that significantly affects the boron (Z=5) scattering factor values. Perhaps, in addition to He, Li and Be, we should have excluded from these graphs the boron data as well. Aside from the issues with boron at high $\sin\theta/\lambda$, the mean $\Delta f_{\%}$ is well below 0.2% at all $\sin\theta/\lambda$ values, while the maximum $\Delta f_{\%}$ values never exceed 2% and for most $\sin\theta/\lambda$ grid points are within 1%.

2.4.1.2 K (Z=19) – Rn (Z=86) / Cf (Z=98)

When calculating the mean Δf and $\Delta f_{\%}$ values at a given $\sin\theta/\lambda$ grid point, the same group of atoms must be selected. However, the DT/CW data extend to Cf (Z=98), RRG data – to U (Z=92), and the SCM atom coverage terminates at Rn (Z=86). As such, while the graphs of Δf and $\Delta f_{\%}$ as a function of the atomic number Z (Figures 2b, 3b, 4b and 5b) extend to Cf (Z=98) (XRSFs for all atoms were calculated for the same $\sin\theta/\lambda$ points, so the averaging of Δf and

$\Delta f_{\%}$ over $\sin\theta/\lambda$ values for each atom is not an issue), for the graphs that show the same parameters plotted as a function of $\sin\theta/\lambda$ where the averaging over the same atoms is desired (Figures 7b and 9b), we used all atoms between K (Z=19) and Rn (Z=86) with the exception of Tc (Z=43) and Pm (Z=61) as the values for those species were not given in the RRG study. The Wang et al. (1996) data do not extend beyond Ar (Z=18) and thus were not included in any of these graphs. The individual plots of Δf as a function of $\sin\theta/\lambda$ for this group of atoms are given in Figures S1-19 – S1-98.

As seen in Figures 2b and 3b, the maximum and mean differences Δf for these atoms are significantly higher than those for the lighter elements. Both graphs show that the deviations in general increase with the atomic number Z. The largest maximum Δf differences not shown in Figure 2b (which was done in order to resolve the fine structure of the plot) include -1.34, -0.62 and -0.44 electrons from the SCM study for Gd (Z=64), Ce (Z=58), and Bi (Z=83), respectively, and +0.33, +0.28 and +0.26 electrons from the CW work for Ho (Z=67), Cf (Z=98) and Bk (Z=97), respectively. The large discrepancies in the SCM study are observed at $\sin\theta/\lambda \leq 2.5$ Å^{-1} for Gd (Figure S1-64), at $\sin\theta/\lambda < 5$ Å^{-1} for Bi (Figure S1-83), and below 0.6 Å^{-1} for Ce (Figure S1-58). The issues in the SCM study with Gd may be related to the numerical integration as the scattering factor evaluated at $\sin\theta/\lambda = 0$ is 63.99997 instead of the expected exact value of 64. However, it does not explain the shape of the Δf vs $\sin\theta/\lambda$ curve shown in Figure S1-64. The largest deviations in the CW work, occur below 0.5 Å^{-1} for Ho (Figure S1-67), and below 0.3 Å^{-1} and above 2 Å^{-1} for Bk (Figure S1-97) and Cf (Figure S1-98). The large discrepancies in the CW data at high $\sin\theta/\lambda$ (> 2 Å^{-1}) that we observe for Am, Cm, Bk and Cf (Figures S1-95 – S1-98) are likely due to issues with extrapolations of XRSFs (Fox, O'Keefe & Tabbernor, 1989) because for all these (and other heavy atoms, such Fr, Ra, Ac, Np and Pu) there is an abrupt and inconsistent increase in Δf above 2 Å^{-1} (see also Figures 10 and 11).

In general, Figures 2b and 3b show a very good agreement for many atoms between our results and those from the other three studies. However, it is surprising to see how the SCM data show the smallest Δf deviations for some atoms and the largest Δf deviations for others. A similar trend is observed for the DT/CW data but without the extremes. The RRG results display the most consistent behaviour: the mean Δf deviations are never too low and never too high

In terms of the percentile deviations $\Delta f_{\%}$ plotted as a function of *Z* (Figures 4b and 5b), the largest maximum $\Delta f_{\%}$ deviations are found in the SCM study for Gd (-4%), Ce (-1.3%), and Bi (-1.2%). It is not unexpected because the very same species showed the largest Δf deviations. For the CW, study there is a surprising maximum $\Delta f_{\%}$ value of 2.1% for Nb (*Z*=41), a number of $\Delta f_{\%}$ values between -1.1 and -1.17% for a group of atoms between Tb and Hf (*Z* = 65 – 72), and $1 < \Delta f_{\%} < 3\%$ for the atoms with *Z* > 94 (Am, Cm, Bk and Cf). As mentioned below, the RRG data show the most consistent behaviour as the maximum $\Delta f_{\%}$ values almost never exceed 0.5%: the only three exceptions are Ca, Sc and Ti (*Z* = 20 – 22) in which the maximum $\Delta f_{\%}$ values are still under 0.6%.

The graph of mean $\Delta f_{\%}$ plotted as a function of *Z* (Figure 5b) shows a trend that was already seen in Figure 3b. The mean $\Delta f_{\%}$ in the RRG study is consistently between 0.03% and 0.06%. The SCM data display the lowest $\Delta f_{\%}$ values of below 0.01% for some atoms and the highest $\Delta f_{\%}$ for others. The DT $\Delta f_{\%}$ quantities are consistently below those from the RRG study but never go as low as the SCM $\Delta f_{\%}$ values. The mean $\Delta f_{\%}$ discrepancies for the CW study for most atoms are above those from the RRG work.

The maximum and mean Δf are $\Delta f_{\%}$ quantities are plotted as a function of $\sin\theta\,/\lambda$ in Figures 6b, 7b, 8b and 9b. As mentioned above, the mean Δf was calculated by averaging over all atoms between K (*Z*=19) and Rn (*Z*=86) excluding Tc (*Z*=43) and Pm (*Z*=61) as the data

for these atoms were not calculated in the RRG study. The largest maximum and mean Δf deviations are observed in the SCM study below approximately 1 Å^{-1} (Figures 6b and 7b), and the largest Δf deviations are observed for Gd. The DT/CW (for consistency, we had no choice but to combine the data from these two studies as they are mutually exclusive) and RRG maximum and mean Δf values are very similar to each other and significantly smaller at $\sin\theta/\lambda < 0.8$ Å^{-1} than those from the SCM study. However, at $\sin\theta/\lambda > 2.0$ Å^{-1} the maximum and mean Δf deviations from the SCM work become comparable to and even smaller than those for the DT/CW and RRG studies.

The highest $\Delta f_{\%}$ deviations (Figure 8b) are observed for the SCM below 2.5 Å^{-1} and may reach up to 3.5% between 0.4 and 0.8 Å^{-1} in Gd. The largest DT/CW $\Delta f_{\%}$ values of around 2% are found at 2.5 Å^{-1} for Ho and at 4 Å^{-1} for Nb. The largest maximum $\Delta f_{\%}$ deviations for the RRG study are always relatively small, below 0.6%. The situation is somewhat different for the mean $\Delta f_{\%}$ deviations (Figure 9b). The RRG mean $\Delta f_{\%}$ deviations are small (<0.3%) and for $\sin\theta/\lambda \leq 3.0$ Å^{-1} are always below 0.1%. However, while below 1 Å^{-1} the mean DT/CW and RRG $\Delta f_{\%}$ deviations are consistently lower than those from the SCM study, at $\sin\theta/\lambda \geq 2.5$ Å^{-1} the situation is reversed, i.e. the SCM $\Delta f_{\%}$ deviations become the lowest. Indeed, Figure 9b clearly shows that the mean $\Delta f_{\%}$ deviations above 2 Å^{-1} increase in the DT/CW and RRG studies but decrease in the SCM work. It also is somewhat troubling to see an unusually high mean $\Delta f_{\%}$ of 0.27% at 2.5 Å^{-1} in the DT/CW data and 0.24% at 5 Å^{-1} in the RRG values. The former issue has been traced to relatively large deviations at 2.5 Å^{-1} for a number of atoms in the CW data (Table 1). These appear to be outliers as none of the other three studies (this work, RRG and SCM) show a similar behaviour. The problem is likely related to some issues with extrapolation of XRSFs in the study by Fox, O'Keefe & Tabbernor (1989).

2.4.1.3 Fr (Z=87) – Cf (Z=98)

In Figure 10a we show both the maximum and mean Δf deviations for the atoms between Fr (Z=87) and U (Z=92) plotted as a function of $\sin\theta/\lambda$, while Figure 10b shows the maximum and mean $\Delta f_{\%}$ quantities for the same group of atoms. Because the SCM data terminate at Rn (Z=86), only the DT/CW and RRG values are included. Surprisingly, for these atoms the maximum and mean Δf deviations are very similar in magnitude. At $0.2 \leq s \leq 3.5$ Å^{-1}, the mean Δf deviations in the DT/CW study are consistently lower than 0.01 electrons and below those from the RRG work, while at the other $\sin\theta/\lambda$ grid points they are between 0.01 and 0.02 electrons and essentially the same as in the RRG data. The maximum Δf deviations may reach approximately 0.04 electrons at $\sin\theta/\lambda < 0.3$ Å^{-1} for both studies (almost exclusively, for U), but there is an expected increase of the maximum Δf in the DT/CW data to 0.056 electrons for Ac at 5 Å^{-1}. The latter is likely to an extrapolation issue in Fox, O'Keefe & Tabbernor (1989): the reported value is 5.933 electrons while the numbers from this and RRG studies are 5.989 and 6.000 electrons, respectively (Table 1).

In terms of the maximum and mean percentile deviations ($\Delta f_{\%}$) (Figure 10b), the values for both the DT/CW and RRG studies remain below 0.05% up to approximately 2.5 Å^{-1}, beyond which the agreement starts to deteriorate. The mean $\Delta f_{\%}$ deviations increase steadily up to 0.2% at 5 Å^{-1} but then drop to below 0.06% at 6 Å^{-1}. The maximum $\Delta f_{\%}$ deviation in the DT/CW study reaches 0.94% at 5 Å^{-1} for Ac (which is likely due to an extrapolation issue discussed above), while that in the RRG work maxes out at -0.22% for Th, also at 5 Å^{-1}.

Finally, in Figure 11 we present results of the same analysis for the elements with Z = 93 – 98 (Np, Pu, Am, Cm, Bk and Cf) but this time using only the CW scattering factors with the Fox, O'Keefe & Tabbernor (1989) extrapolations above 2 Å^{-1} (the RRG and DT data terminate at U). Not unexpectedly, the largest maximum Δf deviations are observed for Cf (the heaviest element in the group) and reach 0.28 electrons at 0.15 Å^{-1} and -0.22 electrons at

3 and 3.5 Å^{-1} (Figure 11a). It is interesting to note how the agreement between our values and CW data consistently improves between 0.4 and 2 Å^{-1} but immediately starts to deteriorate in the $\sin\theta/\lambda$ region of the extrapolated values (Fox, O'Keefe & Tabbernor, 1989).

The maximum and mean percentile deviations ($\Delta f_{\%}$) (Figure 11b), stay more or less constant and well below 0.5% up to $\sin\theta/\lambda = 2$ Å^{-1} but then both start to increase abruptly and randomly, which is very likely due to extrapolation issues. The same effect is clearly observed in Figures S1-93 – S1-98.

Overall, we are very pleased to see the extent to which our new XRSFs compare with the previously determined values (Doyle & Turner, 1968; Cromer & Waber, 1968; Rez, Rez & Grant, 1994; Wang et al., 1996; Su & Coppens, 1997 / Macchi & Coppens, 2001; Maslen, Fox & O'Keefe, 2006). The differences that we observe are likely due to different approaches such as a single-configuration vs multiconfiguration Dirac-Hartree-Fock, OL/EOL/AL/EAL schemes, nuclear models and the extent of the Breit and QED corrections, as well as precision of the integration of X-ray scattering factors. Regarding the latter, the B-spline representation of one-electron orbitals used in DBSR_HF (Zatsarinny & Froese Fischer, 2016) allows for a very precise radial integration of XRSFs which may not be the case in all the earlier studies, especially when extrapolations were involved.

A summary of possible typos and inconsistencies identified in Maslen, Fox & O'Keefe (2006) are listed in Table 1. We note that because for this analysis we employed the Doyle & Turner (1968) $\sin\theta/\lambda$ grid, which has fewer points than that used in Maslen, Fox & O'Keefe (2006), the list may not be complete. However, Table S3 should contain the correct X-ray scattering factors for all the $\sin\theta/\lambda$ values listed in Maslen, Fox & O'Keefe (2006).

2.4.2 Analytical interpolation in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range

2.4.2.1 Four-Gaussian expansion

Results of the fitting of the calculated X-ray scattering factors in the $0 \leq s \leq 2$ Å^{-1} range using equation (39) with $m = 4$ are given in Tables S4 and S5. The latter includes the

comparison of our results in terms of the fitted parameters and the quality of the fit with the benchmark values of Doyle & Turner (1968), Cromer & Waber (1968), and Maslen, Fox & O'Keefe (2006). We note that in both tables, our statistics of the fit are reported with one extra digit as compared to the previously published values.

Because the starting parameters for fitting in our work were taken from the previous studies (Doyle & Turner, 1968; Cromer & Waber, 1968; Maslen, Fox & O'Keefe, 2006), our optimized values are not too different from the original data. The comparison of the two sets of fits for atoms with Z = 2 – 98 in terms of the differences in the maximum and mean errors of the interpolating function is presented in Table S5 and conveniently summarized in Figure 12a. The negative quantities identify a better fit achieved in this work. Note that for Es (Z=99) – Og (Z=118), we started with the coefficients for Cf (Z=98) from Cromer & Waber (1968) since it seems that no interpolation coefficients are available in literature for these atoms, which is why they are not included in the comparison presented in Table S5 and Figure 12a.

The analysis of Figure 12a shows that while the maximum error of the interpolating function (●) in our study decreases for some atoms, it increases for some others. Especially disturbing are the noticeable increases in the maximum error for Tl (Z=81) and Pb (Z=82) by 0.01 and 0.0075 electrons, respectively. An exhaustive search in the parameter space for a lower minimum of the error function for these two atoms was not successful. It seems that the despite using a more sophisticated fitting approach, the differences in the numerical values of XRSFs relative to the works of Doyle & Turner (1968) and Cromer & Waber (1968) cause decline in the quality of the fit in our study for a number of atoms.

A similar behaviour is observed for the mean error (Figure 12a, Δ) though the increases for Tl and Pb are not as drastic – the mean error for these atoms increases by 0.0025 and 0.0019 electrons, respectively. The statistical parameter E defined by Doyle & Turner (1968) as the

"*root mean square value of the deviation δ_i between theoretical and fitted f-values expressed as a percentage of $f(0)$*":

$$E = \frac{100}{f(0)}\left[\frac{\sum_{i=1}^{N_f}\delta_i^2}{N_f}\right]^{1/2} \tag{43}$$

where N_f is the number of the scattering factors that is equal to 201 in Doyle & Turner (1968) and this work, shows that indeed the quality of the fit for Pb (Z=82) is slightly worse in this work ($E = 0.0324$) than in Doyle & Turner (1968) ($E = 0.0296$). A similar situation is observed for Bi (Z=83) for which E in our study (0.0289) is higher than that (0.0279) in Doyle & Turner (1968). However, for some other heavy atoms, such as Hg (Z=80), Rn (Z=86) and U (Z=92), the parameter E is essentially identical in the two studies: 0.0247 and 0.0247 for Hg, 0.0166 and 0.0167 for Rn, and 0.0139 and 0.0143 for U for our work and the Doyle & Turner (1968) study, respectively. It is interesting to note that for Ho (Z=67) the mean error in our work is higher by 0.0013 while the maximum error is lower by 0.0015 electrons.

In summary, out of a total of 194 maximum and mean errors, in 104 instances the error is lower in in our study. The mean negative difference (a better fit in this work), is -0.0005 electrons while the mean positive difference is +0.0007 electrons. As mentioned above, the fact that we do not always get a better fit despite using more advanced computing methods is attributed to the differences in the calculated scattering factors relative to the previous studies.

2.4.2.2 Five-Gaussian expansion

Extending the expansion (39) to $m = 5$ by including an additional Gaussian function, $a_5 \exp(-b_5 \sin\theta^2/\lambda^2)$, significantly lowers the maximum and mean errors of the interpolating function for an overwhelming majority of the atoms (Tables S6 and S7, Figure 12b). For this fit, the starting parameters were taken from our optimized parameters with $m = 4$ (Table S4) and thus the comparison includes all atoms with Z = 2 – 118. We note that in the fitting procedure for $m = 5$, only those optimization methods that allow boundary constraints

on the variables (DMN*x*B, NL2SOL and NL2SO; see Appendix A) were used. The optimized parameters were constrained between the values of –300 and +300 and the number of optimization iterations (steps) was limited to 1,000. A series of trial unconstrained optimization runs with unlimited number of iterations did not lead to a meaningful reduction in the error functions but resulted in a few parameters exceeding a magnitude of 500 which may introduce numerical issues when such a parameter is obtained for an exponential coefficient b_i in expansion (39). As such, the interpolating coefficients listed in Table S6 represent a compromise between accuracy and precision.

For lighter elements (Z is approximately below 22), the maximum error (Figure 12b, ●) is reduced only slightly (except for several atoms such as Mg, Al, Ar, Cu – Se for which the improvement is noticeable) while the mean error (Δ) reduction is within 0.1 electrons. This is not unexpected because for lighter elements even a four-Gaussian fit is able to fairly accurately reproduce the dependence of XRSF on $\sin\theta\,/\lambda$.

For heavier elements, the maximum error is reduced (often, significantly) for most atoms. Surprisingly, for K, Mn, Pd and Ce the maximum error is slightly increased while the mean error is somewhat reduced. In fact, for many atoms with $Z < 60$ there is a noticeable improvement in the maximum error while the mean error is reduced only marginally. At this time, we do not have a reasonable explanation for these observations. We did try to explore the parameter space for several of these atoms by starting with a number of different initial guesses but were unable to find a lower minimum for the error function.

The most dramatic improvement in both the mean and maximum errors of the interpolating function is observed for heavy elements with $Z > 60$. For example, for Tl (Z=81), the maximum and mean errors are reduced from 0.069 to 0.013 and from 0.024 to 0.004 electrons, respectively. The same is true for Pb (Z=82), for which the maximum and mean errors decrease from 0.068 to 0.012, and from 0.023 to 0.004 electrons, respectively.

Improvements of a similar magnitude are also observed for Bi (Z=83). For Cs (Z=55), Ba (Z=56), Fr (Z=87) and Ra (Z=88) the decreases in the errors are smaller but still quite noticeable. Recall that all these atoms (especially Tl and Pb) showed an increase in the maximum error in our $m = 4$ fit relative to the literature data (Doyle & Turner, 1968; Cromer & Waber, 1968).

In general, we think that the significant improvement in accuracy of the interpolating function when increasing the number of the terms in the expansion (39) from four ($m = 4$) to five ($m = 5$), especially in terms of the mean error, warrants the presence of an extra Gaussian function as compared to the literature data (Doyle & Turner, 1968; Cromer & Waber, 1968; Maslen, Fox & O'Keefe, 2006). The accuracy of the analytical interpolations with $m = 5$ in expansion (39) achieved in this work should be more than adequate for even precise X-ray structural work. We also note that the new parametrizations using expansion (39) with $m = 4$ and $m = 5$ are based on the updated relativistic wavefunctions of a higher quality than those used in some of the previous studies. Finally, using the developed SF code and the calculated DBSR_HF wavefunctions, it is straightforward to extend the expansion (39) to $m \geq 6$ if more accurate interpolations in the $0 \leq s \leq 2$ Å^{-1} interval are desired.

2.4.3 Analytical interpolation in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range

The optimized coefficients for the analytical interpolation of the X-ray scattering factors in the $2 \leq s \leq 6$ Å^{-1} range using expansion (40) with $n = 3$ and $n = 4$ are given in Tables S8 and S10, respectively. The comparison with the literature data (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006) for $n = 3$ is presented in Tables S9 and Figure 13a in terms of the correlation coefficient (C) which we define as

$$C = \sqrt{1 - \frac{\sum_{i=1}^{N_f}\left(f_i^{\text{target}} - f_i^{\text{fit}}\right)^2}{\sum_{i=1}^{N_f}\left(f_i^{\text{target}} - \bar{f}^{\text{target}}\right)^2}} \qquad (44)$$

$$\bar{f}^{\text{target}} = \frac{1}{N_f}\sum_{i=1}^{N_f} f_i^{\text{target}} \tag{45}$$

where f^{target} and f^{fit} are the X-ray scattering factors obtained from the DBSR_HF wavefunction and the analytical interpolation, respectively. The parameter N_f is the number of XRSFs used in the fit, which for the $2 \le s \le 6$ Å^{-1} range was equal to 41 ($\sin\theta/\lambda$ step of 0.1 Å^{-1}). However, Table S9 also lists the maximum and mean errors of the interpolating function (40) with $n = 3$ obtained in this work. The changes in these quantities upon extending the expansion (40) from $n = 3$ to $n = 4$ are illustrated in Figure 14, while the change in the correlation coefficient *C* is shown in Figure 13b.

2.4.3.1 Four-term expansion

The analysis of Table S9 and Figure 13a in which we plot the differences in the correlation coefficient, *C* (a positive value represents higher *C* and thus a better fit in our work), shows that for 54 atoms our fit is somewhat better, for 2 atoms it is essentially the same, and for 41 atoms it is slightly worse than that in Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006). The maximum positive improvement for *C* (ΔC) is 0.0011 for Cf (Z=98), while the maximum negative ΔC of –0.0003 is observed for Cr (Z=24). The mean reduction of *C* is –0.0001 which is not significant because of the rounding issues related to the fact that the benchmark values for *C* in Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006) have been reported to only four decimal digits. The mean correlation coefficient improvement (positive ΔC) of 0.0003 is slightly more substantial than the decrease in *C*.

2.4.3.2 Five-term expansion

When extending the expansion (40) to five terms ($n = 4$) the improvement in the quality of the fit is significant as shown in Tables S10 and Figures 13b and 14. Note that because we compare the results of fits with $n = 4$ with our own data for $n = 3$, we are able to analyse not only the

differences in the correlation coefficient *C* (Figure 13b), but also the changes in the maximum and mean errors of the interpolating function (Figure 14).

Extending expansion (40) to $n = 4$, improves the correlation coefficient *C* for 114 out of 117 atoms (for the remaining three atoms, ΔC=0) with the mean improvement (positive ΔC) of 0.0003. The maximum improvement of 0.00097 is observed for Te (Z=52) in which the correlation coefficient increases from 0.998978 for $n = 3$ to 0.999947 for $n = 4$. Figure 13b also shows a very peculiar periodic trend: the most significant improvements are observed for the atoms with *Z* around 21, 33, 52, 73, 99, forming five well-defined maxima around these atoms, while in between the maxima (Z < 15, $Z \approx$ 27, 41, 60, 82 and 110) there is very little to no improvement.

In Figure 14 we show the maximum (Figure 14a) and mean (Figure 14b) errors of the interpolating function (40) with $n = 3$ (●) and $n = 4$ (○), respectively (see also Tables S9 and S11). For both expansions the errors increase with the atomic number *Z* but not linearly – there is a "sinusoidal" trend similar to that observed for ΔC in Figure 13b. It seems that the form of the expansion (40) works well for certain atoms (with *Z* around 20, 32, 51, 70 and 95 and 118) and increasing the number of terms from three to four significantly improves the quality of the fit, which is clearly reflected in all the statistical parameters. For example, the maximum errors decrease by as much as 0.2 – 0.3 electrons for the elements with $90 \leq Z \leq 105$ (Th – Db). Even for lighter elements such as for example, those between Ag and Cs (Z = 47 – 55) the maximum error is reduced by 0.10 – 0.14 electrons. However, for the atoms with *Z* around 25, 40, 60, 82 and 110, the form of the expansion (40) does not seem to be ideal and thus increasing the level of the expansion does not lead to a better fit.

Because expansion (40) is very simple and computationally not expensive, the presented parametrization with $n = 4$ (Table S10) is clearly a preferred option over that for

$n = 3$ (Table S8) as for the overwhelming majority of the atoms it provides a more accurate interpolation (higher correlation coefficient, and lower maximum and mean errors). If one desires even higher accuracy of the interpolating functions between 2 and 6 $Å^{-1}$, the SF code can be used to model the extended function (40) with $n \geq 4$ or function (39) with $n \geq 5$.

2.5 Summary and Concluding remarks

An updated set of the X-ray scattering factors (XRSFs) for neutral atoms (elements) with Z = 2 (He) – 118 (Og) have been determined from the fully relativistic Dirac-Hartree-Fock wavefunctions calculated in the recently developed computer program DBSR_HF (Zatsarinny & Froese Fischer, 2016) in which the radial components of the relativistic orbitals are represented via a B-spline basis. The non-relativistic electron ground state configurations of the atoms (Martin, 2021; Guerra et al., 2017; WebElements, 2022) were automatically expanded in DBSR_HF into suitable sets of relativistic configuration state function (CSFs) with appropriate statistical weights which makes the used approach similar to the average level (AL) and extended average level (EAL) formalisms originally developed by Grant, Mayers & Pyper (1976) and available in GRASP (Grant et al., 1980; Dyall et al., 1989). In comparison to the relativistic wavefunctions of Coulthard (1967) and Mann (1968) that provided a source for the currently recommended XRSFs (Doyle & Turner, 1968; Cromer & Waber, 1968; Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe , 2006) and are limited to Z = 98 (Cf), the DBSR_HF calculations were performed for all elements with Z = 2 – 118, used a dense B-spline radial grid for the one electron orbital wavefunctions, included the Breit interaction correction, and employed the Fermi distribution function for the description of a nuclear charge density. A detailed comparison of the generated relativistic wavefunctions in terms of energies, and local and integrated charge density properties with those from a number of previous studies (Rez, Rez & Grant, 1994; Visscher & Dyall, 1997; Guerra et al., 2017; Tatewaki, Yamamoto & Hatano, 2017) confirmed the quality of the calculations.

The X-ray scattering factors were generated in a newly developed Fortran program SF interfaced to DBSR_HF (Zatsarinny & Froese Fischer, 2016). A combination of the B-spline basis and a fine radial grid employed in the calculations allowed for an integration of XRSFs to a precision of at least eight decimal digits though in the final set of tables we have rounded off the numerical values to five decimals.

The comparison of the redetermined X-ray scattering factors with those listed in the 2006 edition of volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) revealed a number of possible typos and inconsistencies in the published data that we have summarized in Table 1 of this manuscript. However, because the analysis was based on a "truncated" $\sin\theta/\lambda$ grid of Doyle & Turner (1968), the list may not be complete. That said, the new X-ray scattering factors included in Table S3 cover all $\sin\theta/\lambda$ values listed in Maslen, Fox & O'Keefe (2006).

The X-ray scattering factors obtained in this study show a very reasonable agreement with those determined by Doyle & Turner (1968) using relativistic wavefunctions of Coulthard (1967), and Rez, Rez & Grant (1994) who used the multiconfiguration Dirac-Hartree-Fock in the extended average level (EAL) mode (Grant et al., 1980; Dyall et al., 1989). When compared with the values reported by Wang, Smith, Bunge & Jáiregui (1996) (that unfortunately are limited to He – Ar) and Su & Coppens (1997) / Macchi & Coppens (2001) (that terminate at Z = 86, Rn), the agreement is excellent for the majority of atoms at high $\sin\theta/\lambda$ values (above ~0.8 Å^{-1}), but for atoms located in the middle of a block of the Periodic Table and at $\sin\theta/\lambda$ below 0.8 Å^{-1}, the discrepancies may be quite significant (which is especially evident for Ce, Gd, and Tl). It is interesting to note that the agreement between our scattering factors and those from Rez, Rez & Grant (1994) show the opposite trend: the agreement is usually better for atoms in the middle of a block and deteriorates toward the periphery. The agreement with the data of Maslen, Fox & O'Keefe (2006) worsens for the

atoms for which XRSFs were derived by Cromer & Waber (1968) using the wavefunctions of Mann (1968). It is a little strange as the Mann (1968) calculations included a more advanced nuclear charge density model than that used in the Coulthard (1967) wavefunctions of Doyle & Turner (1968). The disagreement increases at $\sin\theta/\lambda > 2$ Å^{-1} because the original Cromer & Waber (1968) scattering factors terminate at 2 Å^{-1} and the values above 2 Å^{-1} had to be obtained by Fox, O'Keefe & Tabbernor (1989) via extrapolation. In general, whenever significant differences are observed among the X-ray scattering factors from Doyle & Turner (1968) / Cromer & Waber (1968), Rez, Rez & Grant (1994), Wang, Smith, Bunge & Jáiregui (1996) and/or Su & Coppens (1997) / Macchi & Coppens (2001), our values "fall in between" which may be an indication that the derived values represent the best compromise among all the previous studies.

Following the works of Doyle & Turner (1968), Cromer & Waber (1968) and Maslen, Fox & O'Keefe (2006) we have applied a four-Gaussian expansion, equation (39) (Vand, Eiland & Pepinsky, 1957) with $m = 4$, to generate the interpolating functions for the atomic scattering factors in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range. Despite using a more sophisticated non-linear optimization approach and significantly more powerful computer hardware, our four-Gaussian interpolations are, on average, of the same accuracy as the benchmark data. A noticeable increase in the maximum error of the interpolating function for several atoms (Tl, Pb, Bi, Cs, and Ba) relative to the benchmark numbers seems to be a consequence of the differences in the newly generated XRSFs relative to the published values.

Extending the expansion (39) to five Gaussian functions, $m = 5$, leads to a significant decrease in the maximum and mean errors of the interpolating functions for most of the atoms when compared to a four-Gaussian expansion. The improvements are especially pronounced for heavy elements with $Z > 60$. Surprisingly, for several atoms (K, Mn, Pd and Ce) the

decrease in the mean error leads to a tiny (below 0.005 electrons) increase in the maximum error.

For the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range, the atomic X-ray scattering factors were approximated using the function (40) with four terms ($n = 3$) as proposed by Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006). The quality of the analytical approximation was estimated based on the magnitudes of the correlation coefficient (*C*) as well as the maximum and minimum errors of the interpolating function. A comparison with the previous studies in terms of changes in the correlation coefficient ΔC (unfortunately, the two types of errors were not reported in Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006)) showed a noticeable improvement in the correlation coefficient for 54 elements (mean ΔC is 0.0003), for two atoms the fit is essentially of the same quality, and for 41 atoms there is a slight reduction in the correlation coefficient (mean $\Delta C = -0.0001$).

Increasing the number of terms in the expansion (40) to five ($n = 4$) results in a significant improvement of the correlation coefficient for the overwhelming majority of the atoms, including ΔC of 0.0009 – 0.0010 for Sc, Sb, Te, I and Xe. As expected, no improvement in *C* is observed for lighter elements ($Z < 15$) as for these species even the expansion with $n = 3$ is already quite accurate. A somewhat surprising result is a "sinusoidal" trend observed in the plot of ΔC as a function of the atomic number *Z* (Figure 13b), which shows virtually no improvement upon extension of the expansion (40) from $n = 3$ to $n = 4$ for atoms with $Z \approx 27$ (Co), 41 (Nb), 60 (Nd), 82 (Pb) and 110 (Ds). The analysis of the improvements in the maximum and mean interpolation errors (Figure 14) confirms a significant advantage of the five-term interpolating function ($n = 4$) over the four-term expansion ($n = 3$) for all atoms except the ones listed above, for which no improvement in the correlation coefficient was achieved. It is interesting to note how for these atoms ($Z \approx 27$, 41, 60, 82 and 110) the four-

term expansion ($n = 3$) gives the lowest maximum and mean errors relative to other heavy atoms, while it is the opposite in the five-term interpolation ($n = 4$). At this time, we do not have a plausible explanation for this behaviour.

For those who require even higher accuracy of the interpolating functions, the SF code can be easily used to extend the function (39) to $m \geq 6$ in the $0 \leq \sin\theta\,/\lambda \leq 2$ Å^{-1} interval, while for the $2 \leq \sin\theta\,/\lambda \leq 6$ Å^{-1} range, one can either extend the expansion (40) to $n \geq 5$ or use the function (39) with $m \geq 6$.

In summary, we believe that the newly derived relativistic Dirac-Hartree-Fock X-ray scattering factors and the accompanied analytical interpolations using the well-established expansions may be useful in the X-ray diffraction studies. For users who require a higher accuracy, we also present extended expansions that require a very minor modification of the existing crystallographic X-ray diffraction software.

For the second part of the study, we plan to use the described approach to re-determine the X-ray scattering factors for all chemically significant ions. Unlike the currently used XRSFs that for different ions were evaluated at different levels of theory, including the non-relativistic Hartree-Fock and relativistic Dirac-Slater techniques (Maslen, Fox & O'Keefe, 2006), our values will be based on the fully relativistic Dirac-Hartree-Fock method with the inclusion of both the Breit interaction correction and the Fermi distribution function for the description of the nuclear charge density (Zatsarinny & Froese Fischer, 2016).

Acknowledgements We would like to thank the MTSU Chemistry Department, Computational Science Ph.D. Program, and Information Technology Division for their continuous support, and the two anonymous referees for carefully reading the manuscript and providing a few helpful suggestions. A number of symbolic computations for debugging purposes were performed in the *Mathematica* system (Wolfram Research, 2022).

APPENDIX A: Non-linear least squares fitting protocol in the program SF

The fitting procedure in our software SF (Figure 1) uses a straightforward "brute-force" approach. The program includes all the following non-linear least squares minimization options:

1) The Levenberg-Marquardt algorithm is provided via subroutines LMDER and LMDIF adapted from the package MINPACK (Moré, Garbow & Hillstrom, 1981; Moré et al., 1984; Moré, Garbow & Hillstrom, 1999). The former uses the analytical Jacobian matrix while in the latter the Jacobian is calculated via a forward-difference approximation.
2) An unconstrained optimization package UNCMIN (Dennis & Schnabel, 1996; Schnabel, Koontz & Weiss, 1982) can make use of both the analytical gradient and analytical Hessian matrix of the error function though from our experience, it was more beneficial to use the code's functionality to construct the Hessian matrix via the secant update method (Dennis & Schnabel, 1996; Schnabel, Koontz & Weiss, 1982).
3) The adaptive least-squares codes NL2SOL and NL2SNO (Dennis, Gay & Welsch, 1981a, Dennis, Gay & Welsch, 1981b; Gay, 1983) implement a variation on Newton's method "*in which part of the Hessian matrix is computed exactly and part is approximated by a secant (quasi-Newton) updating method*". The NL2SOL requires an analytical Jacobian while NL2SNO computes the Jacobian using a finite difference approximation. A significant advantage of these two codes is the ability to improve convergence from poor starting guesses using a "*model/trust-region technique along with an adaptive choice of the model Hessian*" (Dennis, Gay & Welsch, 1981b).
4) The family of three advanced non-linear optimizers, DMN*F*B, DMN*G*B and DMN*H*B, were taken from the package PORT (Fox, 1984). These are modified versions of the NL2SOL and NL2SNO codes (Dennis, Gay & Welsch, 1981a, Dennis, Gay & Welsch, 1981b; Gay, 1983) that can also apply a simple upper and lower bound constraints on

each variable (Gay, 1981). The subroutine DMN*F*B needs only the minimization function values as the gradient is calculated via finite-differences (forward and central) and the secant approximation is used for construction of the Hessian matrix. DMN*G*B is similar to DMN*F*B but uses an analytical gradient of the minimization function, while DMN*H*B makes use of the analytical gradient and analytical Hessian.

5) Finally, we have also tested the conjugate gradient-based subroutine CGFAM (Gilbert & Nocedal, 1992), a simulated annealing program by Corana et al. (1987), and an updated version of the L-BFGS-B code (Morales & Nocedal, 2011), an older version of which was employed by Su & Coppens (1997, 1998) and Macchi & Coppens (2001), but neither of the codes performed for our minimization problems as well as the ones listed above.

The SF program can call all these subroutines in sequence for a given optimization problem and will automatically select the solution that corresponds to the lowest error function. In most cases, the minimizers LMDER, NL2SOL, UNCMIN and DMN*H*G, and DMN*G*B produced very similar error functions and optimized parameters, but DMN*H*B (that uses the analytical gradient and Hessian) achieved the best solution in the majority of the cases, with the other subroutines performing at about the same level and trailing behind.

The starting parameters for the analytical scattering factors interpolations in the two $\sin\theta/\lambda$ ranges (0–2 Å^{-1} and 2–6 Å^{-1}) were taken from Maslen, Fox & O'Keefe (2006). During optimization in the DMN*x*B and NL2S*xy* families of subroutines, the parameters were constrained to the -300…+300 range. When extending the expansions (39) and (40) to $m = 5$ and $n = 4$, respectively, the starting values for the new parameters (for which no literature values are readily available - Su & Coppens (1997, 1998) and Macchi & Coppens (2001) use six-Gaussian expansions) were generated using the following recipe. First, we optimized

parameters for the $m = 4$ and $n = 3$ expansions. Second, for each new parameter, a series of initial values were created:

a) for expansion (39): $\pm 100, \pm 50, \pm 10, \pm 5, \pm 1, \pm 0.5, \pm 0.1,$

b) for expansion (40): $\pm 1, \pm 0.5, \pm 0.1, \pm 0.05, \pm 0.01, \pm 0.005, \pm 0.001, \pm 0.0005, \pm 0.0001.$

Third, a number of initial guesses were produced that included the already optimized parameters from the $m = 4$ and $n = 3$ expansions *and* all permutations of the generated values for the new parameters. Thus, for the expansion (39), a total of 196 initial sets for the optimized parameters were created (14×14 : two parameters with 14 × initial guesses for each parameter), while that number for the expansion (40) was 18 (a single new parameter with 18 × initial guesses). Finally, all the generated initial guesses were supplied to each of the five optimization subroutines (DMN*H*B, DMN*G*B, DMN*F*B, NL2SOL and NL2SNO) that allow simple bounds on the variables – the unconstrained optimizers were not used in the extended fits. Initial guesses that produced an overflow in the error function were immediately discarded.

While the procedure may seem computationally expensive, it takes up to an hour (depending on the optimization input parameters) on an AMD Ryzen 9 5950X desktop processor to integrate and fit scattering factors for all 117 atoms (Z = 2 – 118), including the $m = 5$ and $n = 4$ expansions.

Appendix B: Assessment of the quality of the calculated wavefunctions

The quality of the wavefunctions generated in DBSR_HF has been accessed by an exhaustive comparison with the previous studies in terms of the following properties:

1) total electronic energies (Wang et al., 1996; Su & Coppens, 1997; Visscher & Dyall, 1997; Tatewaki, Yamamoto & Hatano, 2017),
2) orbital (spinor) energies and mean radii (Tatewaki, Yamamoto & Hatano, 2017),
3) orbital (spinor) charge density maxima radii (Guerra et al., 2017), and
4) atom mean radii and mean spherical radii (Guerra et al., 2017).

For clarity, the comparison for entries 2) and 3) was performed separately for the core and valence orbitals. The expression for the mean radius of the i-th orbital (spinor) $\langle r_i \rangle$ is

$$\langle r_i \rangle = \int_0^\infty r \, [P_i(r)^2 + Q_i(r)^2 \,] \, \mathrm{d}r \tag{46}$$

where $P_i(r)$ and $Q_i(r)$ are the large and the small components of the radial wavefunction of the i-th orbital, respectively. These values are printed out by default in DBSR_HF and for debugging purposes are also recalculated in our SF code (a warning is given in SF when the two values disagree by more than 10^{-6} atomic units). Following (Guerra et al., 2017) we construct the total radial electron density

$$r^2\rho(r) = \sum_i^{N_{\mathrm{occ}}} n_i \, [P_i(r)^2 + Q_i(r)^2 \,] \tag{47}$$

where n_i is the "effective" population of the i-th orbital (spinor) and the summation proceeds over all N_{occ} occupied orbitals, which is normalized as (Guerra et al., 2017)

$$N_{\mathrm{e}} = \int_0^\infty r^2\rho(r) \, \mathrm{d}r \tag{48}$$

where N_{e} is the total number of electrons in an atom. Then, the atomic mean radius $\langle r_{\mathrm{at.}} \rangle$ and the atomic mean spherical radius $\sqrt{\langle r_{\mathrm{at.}}^2 \rangle}$ are calculated as in Guerra et al. (2017):

$$\langle r_{\text{at.}}^{(p)} \rangle = \frac{1}{N_{\text{e}}} \int_0^{\infty} r^{(p)}\, r^2 \rho(r)\, \mathrm{d}r \tag{49}$$

where $p = 1$ for the atomic mean radius, and $p = 2$ for the atomic mean spherical radius.

B1: Total atom energies

The total electronic energies of atoms determined in this study are compared to those from Wang et al. (1996), Su & Coppens (1997), Visscher & Dyall (1997), and Tatewaki, Yamamoto & Hatano (2017) in Table S12 and the summary is given in Table 2. Since different studies include a different number of atoms, the maximum and mean energy differences with respect to the values obtained in this work are examined for several groups of atoms.

For He (Z=2) – Ar (Z=18), the agreement between the energies obtained in this work are in an excellent agreement with those from Visscher & Dyall (1997) and Tatewaki, Yamamoto & Hatano (2017) which is not surprising as the Visscher & Dyall (1997) calculations employed the average level (AL) model (Grant et al., 1980; Dyall et al., 1989) and Tatewaki, Yamamoto & Hatano (2017) used the average-of-configuration (AOC) approximation (Desclaux, 1973). That said, our energies are closer the Visscher & Dyall (1997) values because both calculations used the Fermi charge distribution function for a nuclear charge density (Johnson & Soff, 1985; Zatsarinny & Froese Fischer, 2016) while Tatewaki, Yamamoto & Hatano (2017) employed the uniform nuclear charge distribution model. We note that Visscher & Dyall (1997) also reported the total electronic energies for the uniform sphere model of the nuclear charge density, the same as used by Tatewaki, Yamamoto & Hatano (2017), but since the Fermi model is considered to be more accurate, we did not use the uniform charge distribution model in our calculations. The energies from Wang et al. (1996) and Su & Coppens (1997) show significant deviations from our work and from each other: the mean energy difference between those two studies is 4.0×10^{-2} hartrees and the maximum difference of 0.14 hartrees is observed for Mg (Z=12). While the differences

between our work and these two studies can be explained by a different level of approximation used (EAL vs OL), it is not clear why the two studies that were performed at the same level (OL) show significant discrepancies. In addition, except for C, N, O, and P, the energies of Su & Coppens (1997) are always higher (less negative) than our values, which is somewhat surprising as Su & Coppens (1997) performed multiconfiguration DHF calculations, and just like in our study, included both the Breit and vacuum polarization corrections (our energies also include the electrons self-polarization correction which is usually smaller in magnitude than the Breit correction but larger than the vacuum polarization term).

The discrepancies between our energies and those from Su & Coppens (1997) get much larger for K (Z=19) – Xe (Z=54): the mean energy difference is almost 2 hartrees and the maximum deviation of 5.1 hartree is found for Xe (with our value being more negative). For this group of atoms, the agreement between our work and Visscher & Dyall (1997) and Tatewaki, Yamamoto & Hatano (2017) also starts to slightly deteriorate though the mean energy differences are on the order of 10^{-4} – 10^{-3} hartree, and the maximum deviation is below 0.01 hartree, which are several orders of magnitude smaller than the differences with respect to the values of Su & Coppens (1997).

For Cs (Z=55) – Lr (Z=103), the energy differences between our work and Visscher & Dyall (1997) and Tatewaki, Yamamoto & Hatano (2017) increase to 0.08 and ≈0.7 hartrees for the mean and maximum values, respectively. Surprisingly, for Bk (Z=97) and Lr (Z=103) the discrepancies between our work and Tatewaki, Yamamoto & Hatano (2017) are only 0.120 and 0.042 hartree, respectively, despite the differences in the employed energy configurations (see Table S1 and discussion in section 2.1). The energy difference for Lr (Z=103) between our work and Visscher & Dyall (1997) is about 0.15 hartree despite the differences in the electronic configurations employed in the calculations. Unfortunately, Macchi & Coppens

(2001) who extended the Su & Coppens (1997) calculations to beyond Xe, did not report the total energies for neutral atoms, so no comparison with their data is possible.

Finally, for the heavy elements between Rf (Z=104) and Mt (Z=109) the mean and maximum energy differences between our work and Visscher & Dyall (1997) reach 0.6 and 1.4 hartree (for Mt), respectively, with our energy values being consistently higher. However, considering that the energies for these atoms are on the order of 4×10^4 hartrees, the maximum difference of 1.4 hartree accounts for only about 0.004%.

B2: Orbital energies

Tables S13 and S14 include the comparison of the orbital (spinor) energies between our calculations and the work of Tatewaki, Yamamoto & Hatano (2017) for the core and valence shells, respectively. Because in the previous study the heaviest element calculated was Lr, only the elements with Z = 2 – 103 are included. As mentioned earlier, the spinor parameters from the Visscher & Dyall (1997) calculations were not included in the paper and instead were made available at "http://theochem.chem.rug.nl/~luuk/FiniteNuclei/" that is, unfortunately, no longer accessible.

The average energy difference for a *core* orbital between the two studies calculated over 1334 values is about 0.2 hartree. The largest difference of 8.591 hartree is observed for 1s orbital of Lr (Z=103). In general, the difference in the orbital energy increases with Z and decreases with the quantum numbers n and l. For example, for No (Z=102), the energy difference for 1s is ~8.5 hartree but only 6×10^{-4} hartree for $6p_{3/2}$. In fact, the energy difference for the outermost core orbital is on the order of 0.0005 hartree for essentially all atoms with two notable exceptions: Bk (Z=97) and Lr (Z=103). These atoms show large deviations for all orbitals relative to all other atoms, even for the outermost $6p_{3/2}$ orbital the differences are about 0.1 hartree. However, it is expected considering that different electronic configurations were used in the two studies. The large differences in energies for the innermost orbitals have a

simple origin: the Tatewaki, Yamamoto & Hatano (2017) study employed the uniform nuclear charge distribution model and the Gaussian-type functions, while we used the Fermi-type distribution scheme and the B-spline basis. As such, the observed significant differences in energies for the innermost orbitals that increase with the atomic number is fully expected.

The agreement for the valence orbitals (Table S14) is excellent: the mean energy difference calculated over 359 values is just 0.002 hartree which includes contributions even from Bk (Z=97) and Lr (Z=103). In fact, out of 359 differences, the highest six values that exceed 0.005 hartree belong to the 5f and 7s orbitals of these two atoms. The next eleven largest discrepancies are on the order of 0.001 – 0.002 hartree. When the Bk and Lr energies are excluded from the calculation of statistics, the mean valence orbital energy difference between our work and Tatewaki, Yamamoto & Hatano (2017) is reduced to 0.0004 hartree with the largest difference of 0.002 hartree observed for $4f_{7/2}$ of Rn.

B3: Mean orbital radii

The comparison of the mean orbital radii defined in equation (46) with the values obtained by Tatewaki, Yamamoto & Hatano (2017) is presented in Tables S15 and S16 for the core and valence orbitals, respectively. The agreement for the mean radii of the *core* orbitals in the two studies is excellent (Table S15). The mean difference calculated over all 1334 values, including those for Bk and Lr, is about 0.0001 bohr. There are only four differences above 0.005 bohr: 6s, $6p_{1/2}$ and $6p_{3/2}$ orbitals of Bk, and $6p_{3/2}$ orbital of Lr, which are clearly due to different electronic configurations. Out of the remaining 1330 calculated differences, 1300 are below 0.001 bohr, which is a little surprising considering the extent of differences in the energies for the innermost core orbitals (see discussion above).

The analysis of Table S16 that lists mean radii of the *valence* orbitals shows that the largest discrepancies are observed, as already expected, for Bk (0.27 bohr for 7s, 0.057 for $5f_{7/2}$, 0.046 bohr for $5f_{5/2}$) and Lr (0.19 bohr for 7s). Once these exceptions due to different electronic

configurations are excluded from the calculation of statistics, the average difference for the mean radius of a valence orbital between the two studies drops to about 0.002 bohr. Out of 353 values, only 31 are above 0.005 bohr with the largest difference of 0.044 bohr found for the $6p_{1/2}$ orbital of Tl (Z=81).

B4: Radii of the orbital charge density maxima

Tables S17 and S18 include the comparison for the radii of the orbital charge density maxima. As before the comparison is done separately for the core (Table S17) and valence (Table S18) orbitals. Unlike in the previous sections, we compare our results to only those given in Guerra et al. (2017) because Tatewaki, Yamamoto & Hatano (2017) did not report these parameters. However, the comparison with the Guerra et al. (2017) data includes all elements with Z = 2 – 118.

The agreement between the two studies for the *core* orbital charge density maxima is excellent (Table S17). The mean difference calculated over 1694 entries is only 0.001 bohr with the largest difference of 0.011 bohr being observed for the $6p_{3/2}$ orbital of Pu (Z=94). In fact, that difference is the only one above 0.01 bohr with 44 more ranging between 0.005 and 0.009_9 bohr. The overwhelming majority of the differences (1144 out of calculated 1694) fall below 0.001 bohr.

Surprisingly, the situation is not as good for the *valence* orbitals (Table S18). The mean difference calculated over 436 values is ≈0.01 bohr. The largest difference of 0.45 bohr is observed for the $7p_{3/2}$ orbital of Fl (Z=114). The remaining eleven differences of 0.10 – 0.44 bohr are found for the following orbitals: 4s of Cr (Z=24), 5s of Nb (Z=41) and Mo (Z=42), $6p_{3/2}$ of Pb (Z=82), $6d_{5/2}$ and $6d_{5/2}$ of Pa (Z=91), U (Z=92) and Np (Z=93), and $6d_{5/2}$ of Cm (Z=96). That said, out of 436 calculated differences, 326 were found to be below 0.005 bohr.

At this time, we are attributing these discrepancies a different set of QED corrections included in our calculations with the DBSR_HF code and the Guerra et al. (2017) calculations

in MCDFGME (Desclaux, Mayers & O'Brien, 1971; Desclaux, 1975; Indelicato & Desclaux, 1990). Guerra et al. (2017) clearly state that their calculations included both the self-energy and vacuum polarization "*self consistently*" even though it is noted that those effects on the effective nuclear charges are "*almost negligible*". Moreover, it is not clear from Guerra et al. (2017) whether their calculations included the Breit correction even though it is very likely that it was the case. In addition, it also possible that we misunderstand the statement in Guerra et al. (2017) regarding the "*monoconfiguration mode*" used in MCDFGME (Desclaux, Mayers & O'Brien, 1971; Desclaux, 1975; Indelicato & Desclaux, 1990): if the "*monoconfiguration mode*" in MCDFGME is similar to the optimal level (OL) model in GRASP92 (Grant et al., 1980; Dyall et al., 1989), it would explain the differences with our EAL calculations.

B5: Atomic mean radii and Atomic mean spherical radii

The final set of benchmark values listed in Table S19 include both the atomic mean and atomic mean spherical radii as defined in equation (49). The agreement between DBSR_HF (this study) and MCDFGME (Guerra et al., 2017) is good but not perfect: the average difference between the two sets of parameters calculated over all atoms is 0.0015 bohr for $\langle r_{\text{at.}} \rangle$ and 0.0049 bohr for $\sqrt{\langle r_{\text{at.}}^2 \rangle}$. For $\langle r_{\text{at.}} \rangle$, the largest discrepancy of 0.014 bohr is observed, surprisingly, for N (Z=7) though it amounts for only about 1% of the mean atomic radius of that atom. Two more differences of above 0.01 bohr are 0.013 for Cr (Z=24) and 0.011 bohr for Fl (Z=114). As expected, the differences for $\sqrt{\langle r_{\text{at.}}^2 \rangle}$ are larger than those for $\langle r_{\text{at.}} \rangle$. The most significant discrepancy is 0.054 bohr for Fl (Z=114), and there are seven more $\sqrt{\langle r_{\text{at.}}^2 \rangle}$ differences in the range of 0.01 – 0.04 bohr for N, Cr, Nb, Mo, Pb, Lr and Nh.

It is a little disturbing to find noticeable differences in the two types of the mean atomic radii for lighter elements such as N (Z=7) and Cr (Z=24) between the two sets of relativistic calculations. As mentioned above, we attribute these disagreements to a different treatment of QED corrections in our work and in Guerra et al. (2017), or due to differences between the

"*monoconfiguration mode*" in MCDFGME and EAL method in DSBR_HF. That said, an excellent agreement between our results and those obtained by Visscher & Dyall (1997) and Tatewaki, Yamamoto & Hatano (2017), and a good overall agreement with the Guerra et al. (2017) data instils confidence in the quality of the DBSR_HF wavefunctions used for the determination of the X-ray scattering factors.

Table 3 The mean and maximum differences in total electronic energies E (atomic units) of the elements relative to the values obtained in this work. A negative maximum ΔE means that the energy in this work is lower. For each maximum value, the element for which that value is observed is listed in parentheses.

	Wang et al. (1996)	**Su & Coppens (1997)**	**Visscher & Dyall (1997)**	**Tatewaki, Yamamoto & Hatano (2017)**
		He (Z=2) – Ar (Z=18)		
Mean $\lvert\Delta E\rvert$	1.7×10^{-2}	4.2×10^{-2}	7.4×10^{-6}	4.6×10^{-4}
Maximum ΔE (element)	0.10 (N)	-0.14 (Mg)	4.5×10^{-5} (Cl)	-1.1×10^{-3} (Ar)
		K (Z=19) – Xe (Z=54)		
Mean $\lvert\Delta E\rvert$	—	1.8	6.8×10^{-4}	2.2×10^{-3}
Maximum ΔE (element)	—	-5.1 (Xe)	-4.4×10^{-3} (Te)	-6.9×10^{-3} (Te)
		Cs (Z=55) – Lr (Z=103)		
Mean $\lvert\Delta E\rvert$	—	—	7.6×10^{-2} †	7.8×10^{-2} †
Maximum ΔE (element)	—	—	0.71 (Am) †	0.68 (Am) †
		Rf (Z=104) – Mt (Z=109)		
Mean $\lvert\Delta E\rvert$	—	—	0.62	—
Maximum ΔE (element)	—	—	1.4 (Mt)	—

† For Bk (Z=97), Tatewaki, Yamamoto & Hatano (2017) used the [Rn]$5f^8 6d 7s^2$ electronic state while Visscher & Dyall (1997) and we used [Rn]$5f^9 7s^2$. For Lr (Z=103), Tatewaki, Yamamoto & Hatano (2017) and Visscher & Dyall (1997) used [Rn]$5f^{14} 6d 7s^2$ while we used [Rn]$5f^{14} 7s^2 7p$ as was determined by Desclaux & Fricke (1980).

Table 4 Possible typos and inconsistencies identified in Maslen, Fox & O'Keefe (2006). Because the analysis was based on the Doyle & Turner (1968) $\sin\theta/\lambda$ grid, the list may not be complete.

Atom	**$\sin\theta/\lambda$ (Å^{-1})**	**Reported *f*-value**	**This work**	**DT/CW**	**RRG**	**WSBJ**	**SCM**
Ne	0.60	2.517	2.790	2.790	2.791	2.790	2.790
	3.50	0.331	0.351	0.351	0.351	0.351	0.351
Si	4.00	0.566	0.556	0.556	0.555	0.556	0.556
Nb	4.00	2.405	2.455	—	2.455	—	2.456
In	6.00	1.746	1.764	1.764	1.762	—	1.764
Ce	2.50	7.117	7.139	—	7.144	—	7.139
Pr	2.50	7.333	7.363	—	7.368	—	7.363
Nd	2.50	7.567	7.601	—	7.606	—	7.601
	3.00	5.930	5.956	—	5.960	—	5.956
Pm	2.50	7.817	7.845	—	—	—	7.844
Gd	2.50	8.683	8.610	—	8.617	—	8.597
Tb	2.50	8.983	8.858	—	8.865	—	8.858
Dy	2.50	9.267	9.114	—	9.121	—	9.114
Ho	2.50	9.533	9.370	—	9.376	—	9.370
Er	2.50	9.783	9.623	—	9.629	—	9.622
Tm	2.50	10.033	9.872	—	9.879	—	9.873
Yb	2.50	10.267	10.118	—	10.125	—	10.120
Lu	2.50	10.500	10.367	—	10.374	—	10.369
Hf	2.50	10.733	10.613	—	10.619	—	10.614
Ta	2.50	10.950	10.854	—	10.860	—	10.855
W	2.50	11.167	11.089	—	11.095	—	11.090
Re	2.50	11.383	11.318	—	11.324	—	11.320
Ac	5.00	5.933	5.989	—	6.000	—	—
Pu	3.00	12.656	12.587	—	—	—	—
	3.50	10.895	10.838	—	—	—	—
Am	3.00	12.838	12.731	—	—	—	—
	3.50	11.095	11.004	—	—	—	—
	5.00	6.713	6.795	—	—	—	—

Table 4 (continued).

Atom	$\sin\theta/\lambda$ (Å^{-1})	Reported *f*-value	This work	DT/CW	RRG	WSBJ	SCM
Cm	2.50	14.988	15.046	—	—	—	—
	3.00	13.019	12.875	—	—	—	—
	3.50	11.295	11.168	—	—	—	—
	5.00	6.825	6.945	—	—	—	—
	6.00	5.414	5.477	—	—	—	—
Bk	2.50	15.150	15.239	—	—	—	—
	3.00	13.200	13.016	—	—	—	—
	3.50	11.495	11.323	—	—	—	—
	5.00	6.937	7.096	—	—	—	—
	6.00	5.484	5.572	—	—	—	—
Cf	2.50	15.311	15.442	—	—	—	—
	3.00	13.381	13.157	—	—	—	—
	3.50	11.695	11.476	—	—	—	—
	5.00	7.049	7.250	—	—	—	—
	6.00	5.553	5.671	—	—	—	—

DT - Doyle & Turner (1968)

CW – Cromer & Waber (1968)

RRG - Rez, Rez & Grant (1994)

WSBJ - Wang, Smith, Bunge & Jáiregui (1996)

SCM - Su & Coppens (1997) and Macchi & Coppens (2001)

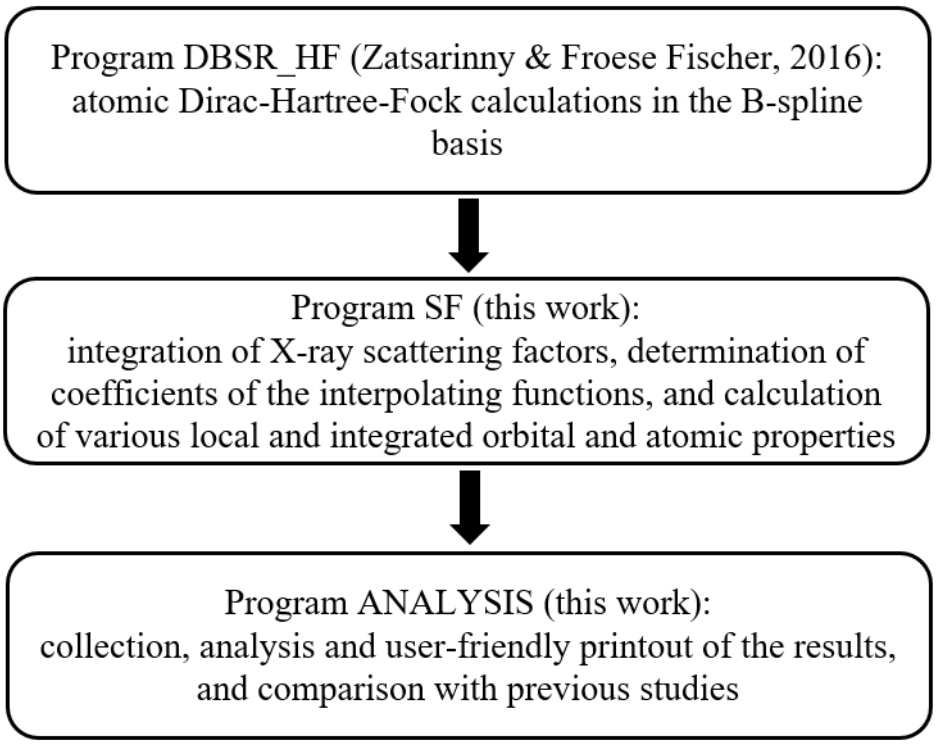


Figure 1 The sequence of steps and the associated software used in the present study.

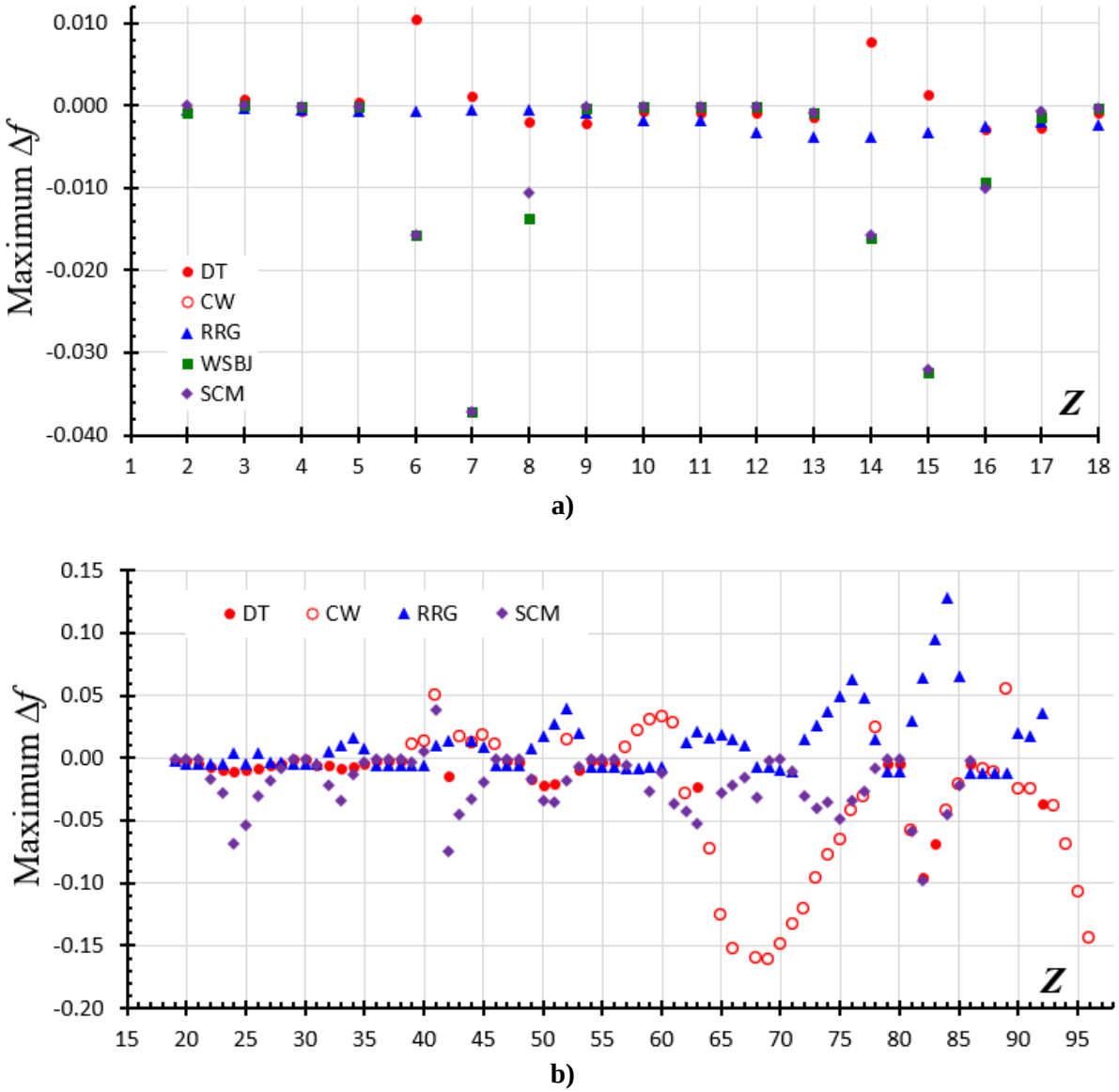


Figure 2 The *maximum* differences in the X-ray scattering factors (Δf) between this work and the previous studies for **a)** He (Z=2) –Ar (Z=18) and **b)** K (Z=19) – Cf (Z=98) plotted as a function of the atomic number, Z. The previous studies are identified as follows: ● DT - Doyle & Turner (1968), ○ CW - Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). The $\sin\theta/\lambda$ grid from the Doyle & Turner (1968) study was used. Not shown in **b)** are the differences of -1.34, -0.62 and -0.44 in the SCM study for Gd (Z=64), Ce (Z=58), and Bi (Z=83), respectively, and +0.33, +0.28 and +0.26 in the CW work for Ho (Z=67), Cf (Z=98) and Bk (Z=97), respectively.

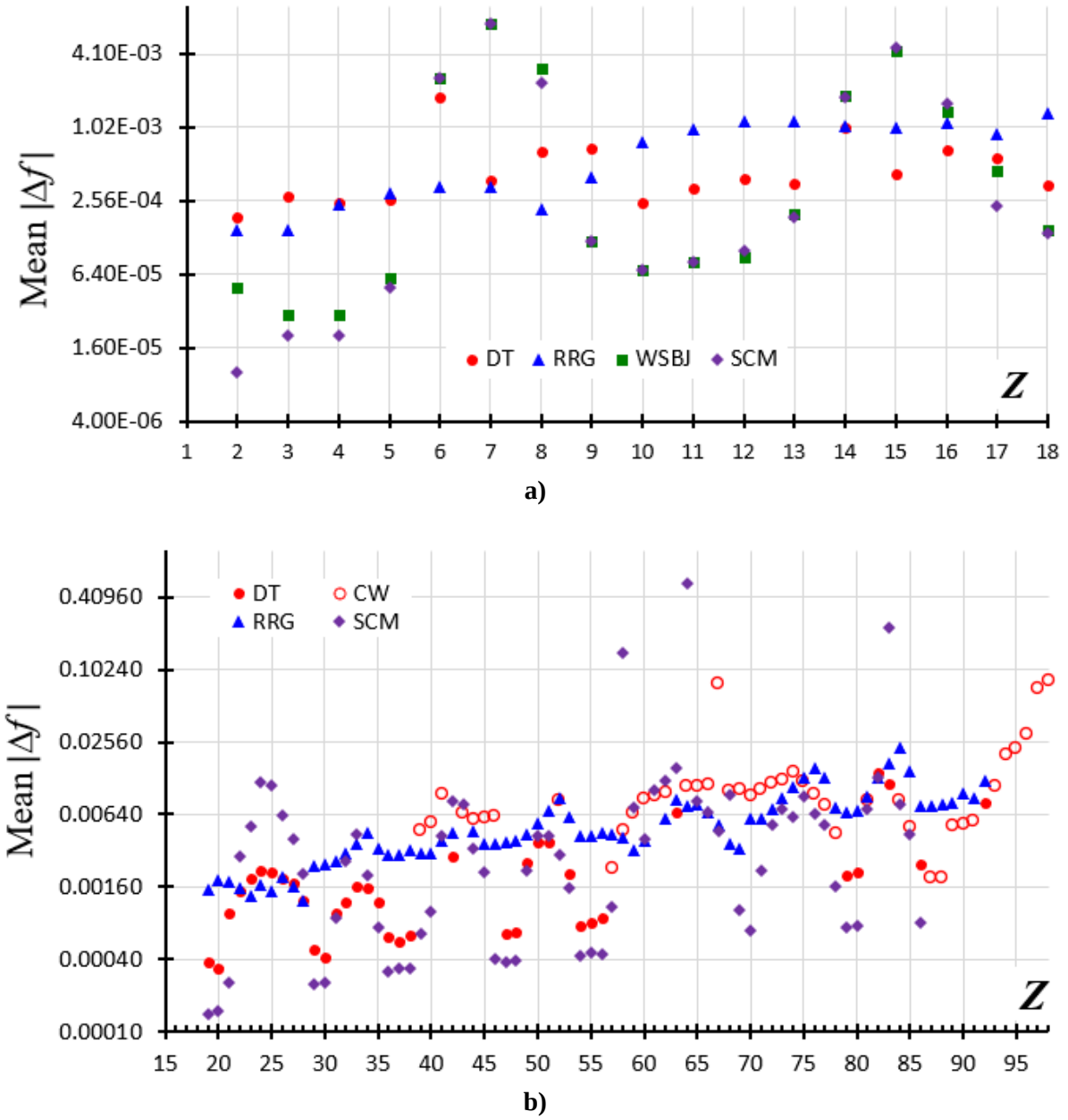
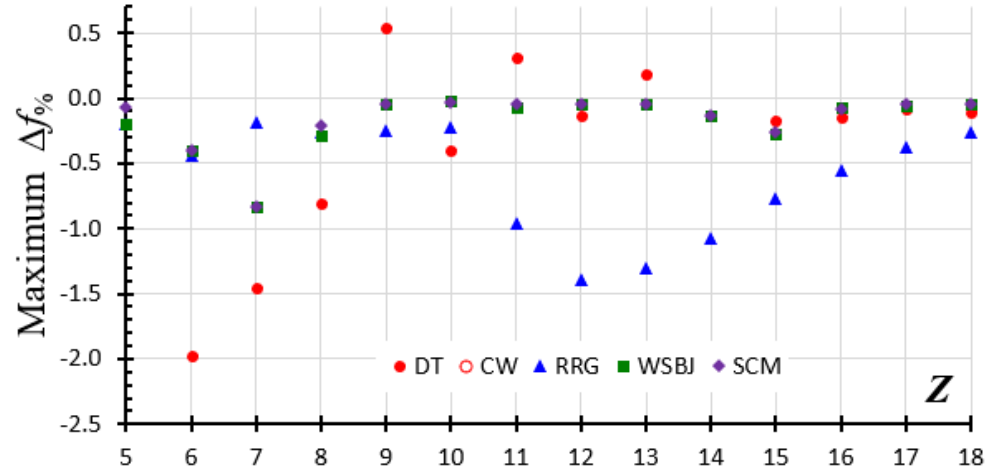
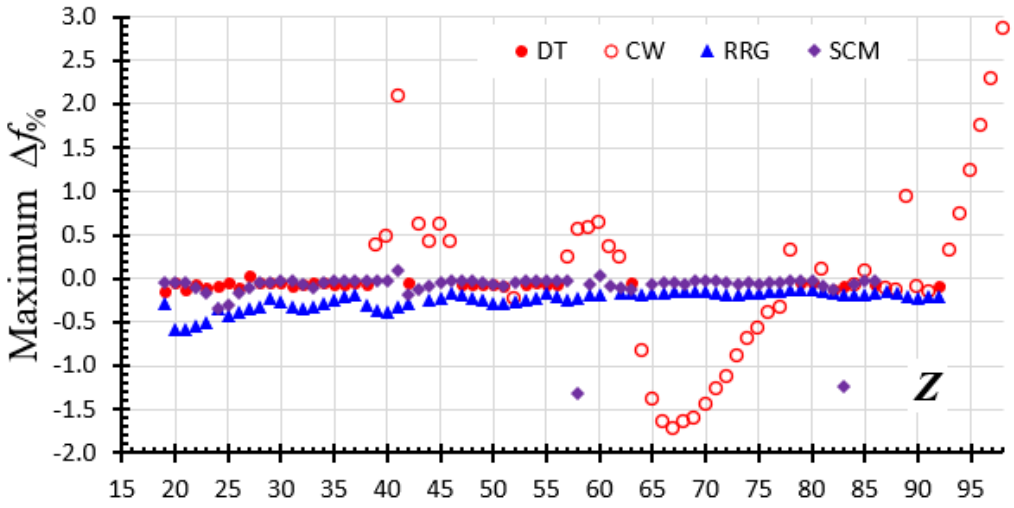


Figure 3 The *mean absolute* differences in the X-ray scattering factors ($|\Delta f|$) between this work and the previous studies for **a)** He (Z=2) – Ar (Z=18) and **b)** K (Z=19) – Cf (Z=98) plotted as a function of the atomic number, Z. The previous studies are identified as follows: ● DT - Doyle & Turner (1968), ○ CW - Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). The averaging was done over all the $\sin\theta/\lambda$ grid points between 0 and 6 Å^{-1} as given in the Doyle & Turner (1968) study. For convenience, a logarithmic base-2 scale is used for the y-axis in both graphs.



a)



b)

Figure 4 . The *maximum percent* differences in the X-ray scattering factors ($\Delta f_{\%}$) between this work and the previous studies for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Cf (Z=98) plotted as a function of the atomic number, Z. The previous studies are identified as follows: ● DT - Doyle & Turner (1968), ○ CW - Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). The grid from the Doyle & Turner (1968) study was used. Not shown in **a)** is the difference of 5% in the DT study for B (Z=5) at 6 Å^{-1} due to the limited number of reported digits (0.00631 vs 0.006), which is also the case for C (0.01274 vs 0.013) and N (0.02267 vs 0.023) at 6 Å^{-1}. Not shown in **b)** is the difference of -4% for Gd (Z=64) in the SCM work.

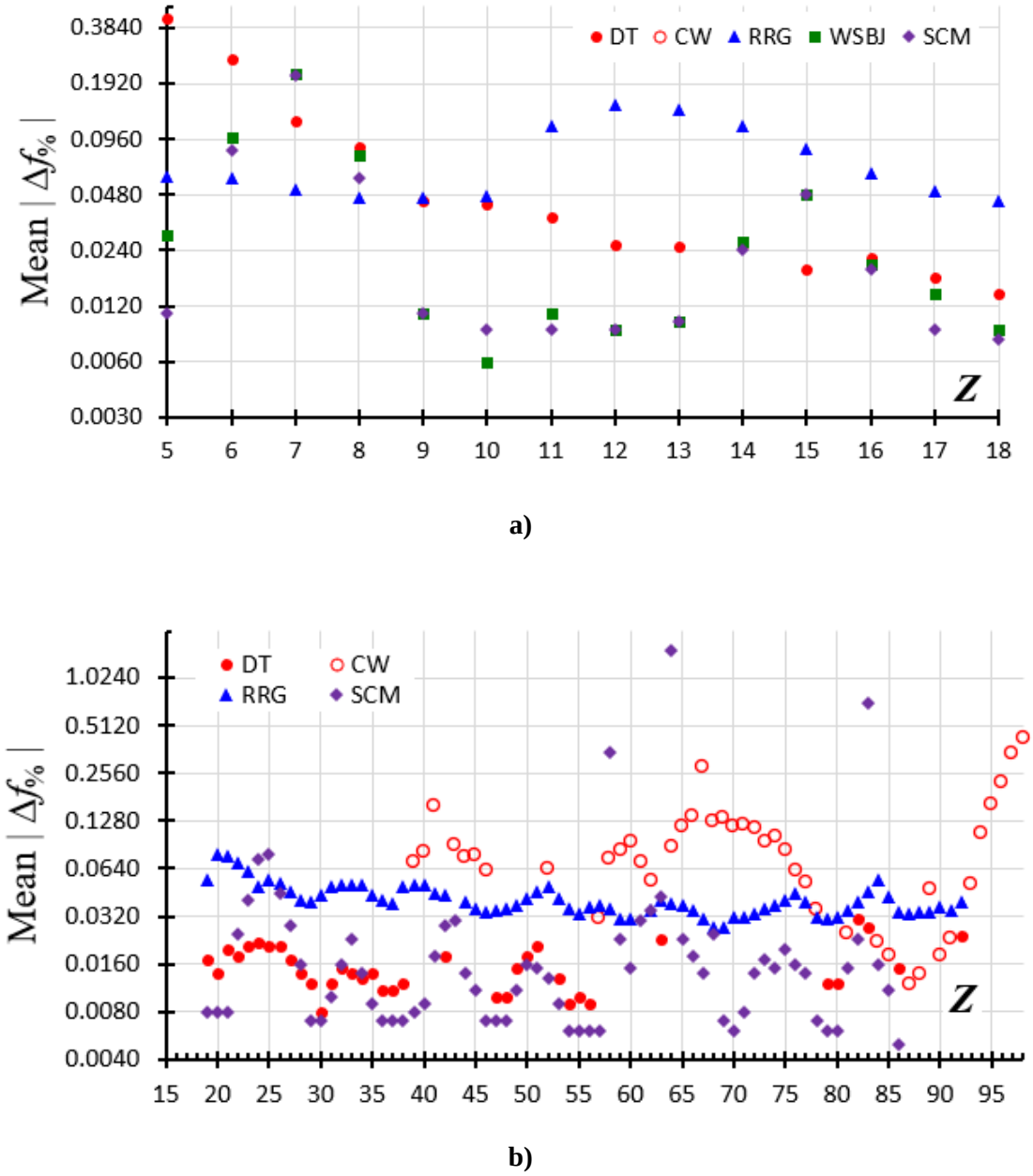
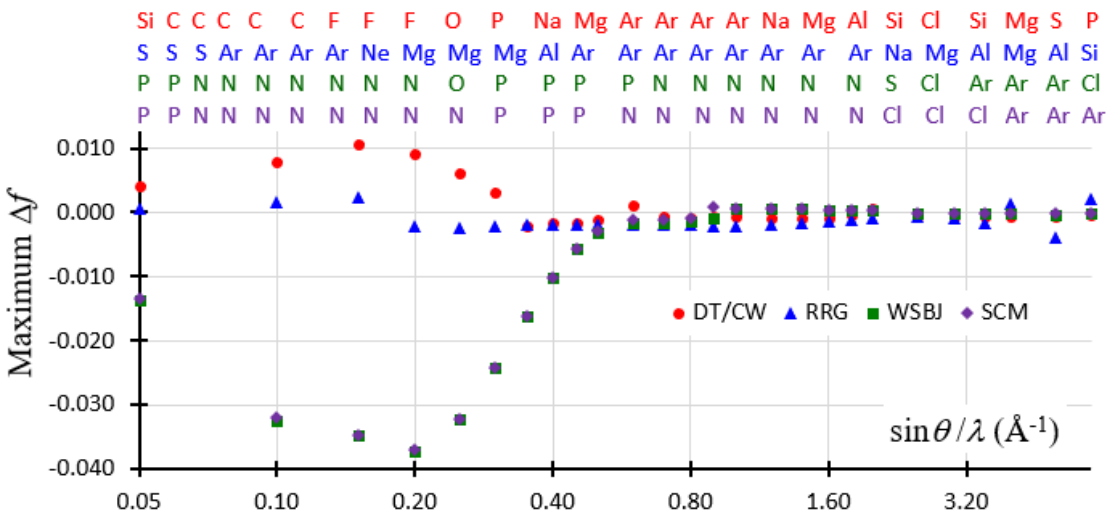
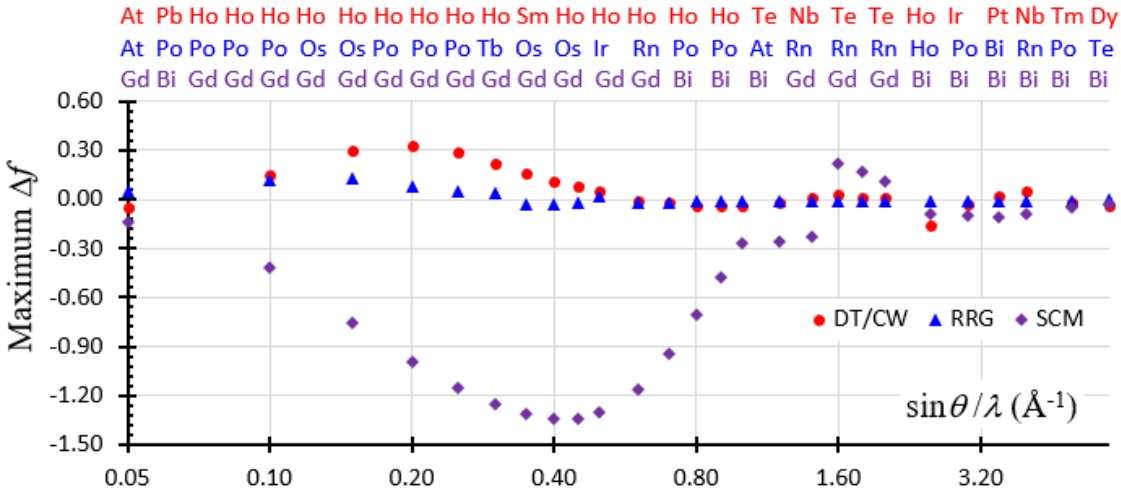


Figure 5 . The *mean absolute percent* differences in the X-ray scattering factors ($|\Delta f_{\%}|$) between this work and the previous studies for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Cf (Z=98) plotted as a function of the atomic number, Z. The previous studies are identified as follows: ● DT - Doyle & Turner (1968), ○ CW - Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). The averaging was done over all the $\sin\theta/\lambda$ grid points between 0 and 6 Å^{-1} as given in the Doyle & Turner (1968) study. For convenience, a logarithmic base-2 scale is used for the y-axis in both graphs.

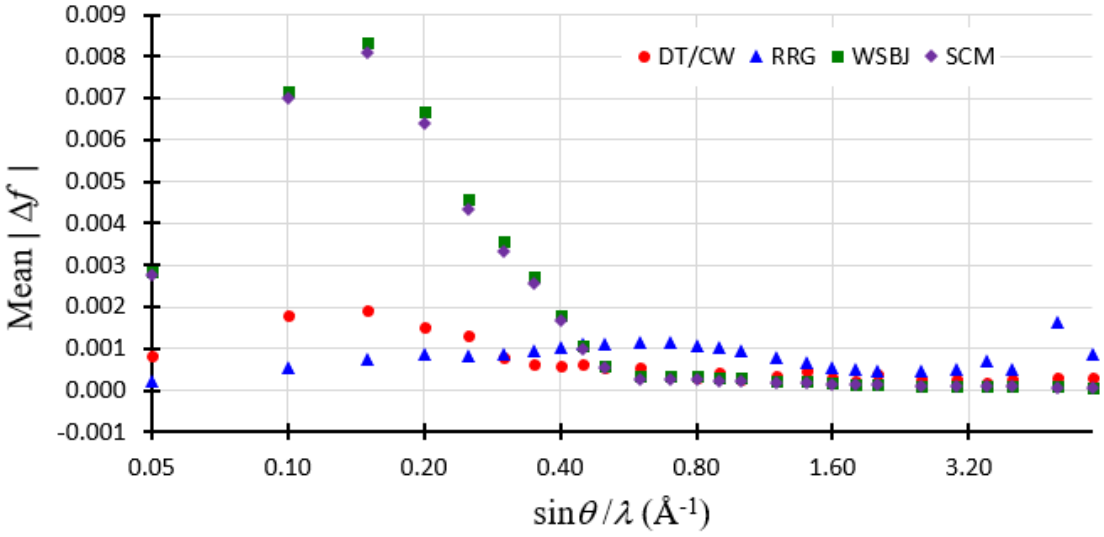


a)

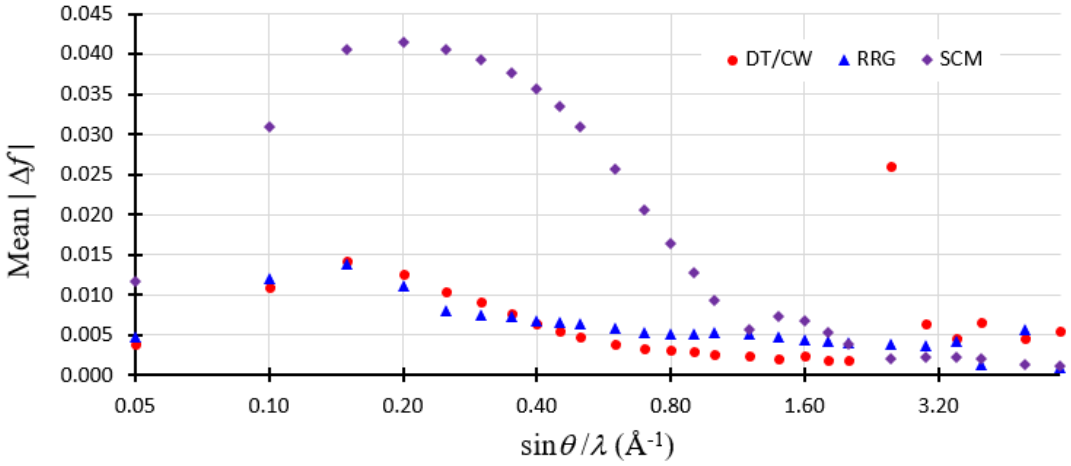


b)

Figure 6 . The *maximum* differences in the X-ray scattering factors (Δf) between this work and the previous studies for each $\sin\theta/\lambda$ grid point in the 0 – 6 Å^{-1} range (Doyle & Turner, 1968) for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Rn (Z=86). The previous studies are identified as follows: ● DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). The elements Tc (Z=43) and Pm (Z=61) were excluded as no data are given in the RRG study for these species. For convenience, a logarithmic base-2 scale is used for the x-axis in both graphs. The symbols above each graph identify elements for which the maximum differences Δf were observed at the $\sin\theta/\lambda$ grid points.

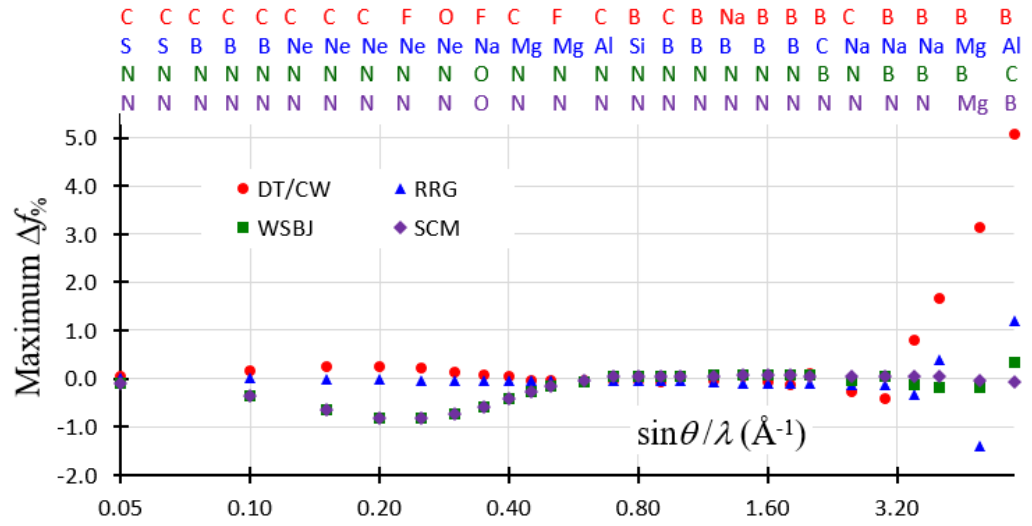


a)

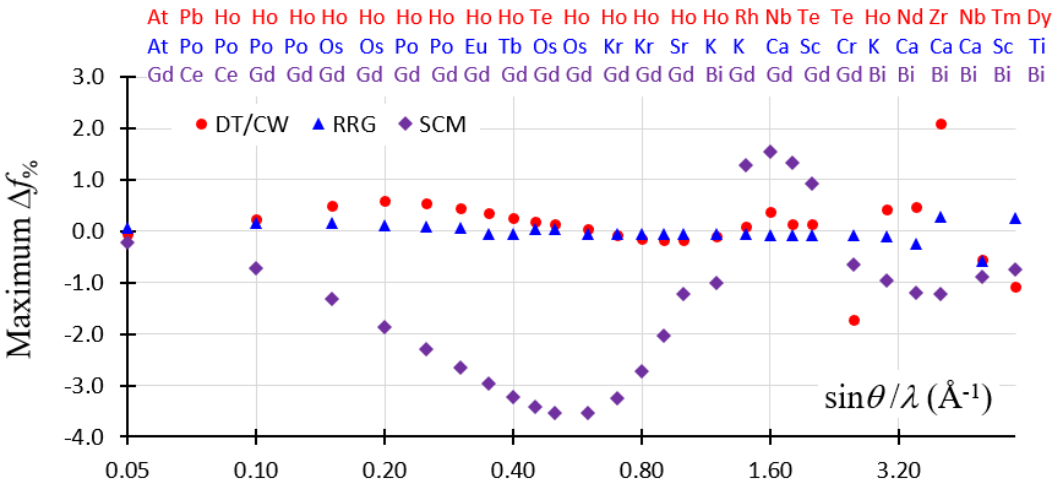


b)

Figure 7 . The *mean absolute* differences in the X-ray scattering factors ($|\Delta f|$) between this work and the previous studies for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Rn (Z=86) plotted as a function of $\sin\theta/\lambda$. The elements Tc (Z=43) and Pm (Z=61) were excluded as no data are given in the RRG study for these species. The averaging at each $\sin\theta/\lambda$ grid point between 0 and 6 Å^{-1} (Doyle & Turner, 1968) was done over a specified group of atoms: **a)** B – Ar, and **b)** K – Rn. The previous studies are identified as follows: ● DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). For convenience, a logarithmic base-2 scale is used for the *x*-axis in both graphs.

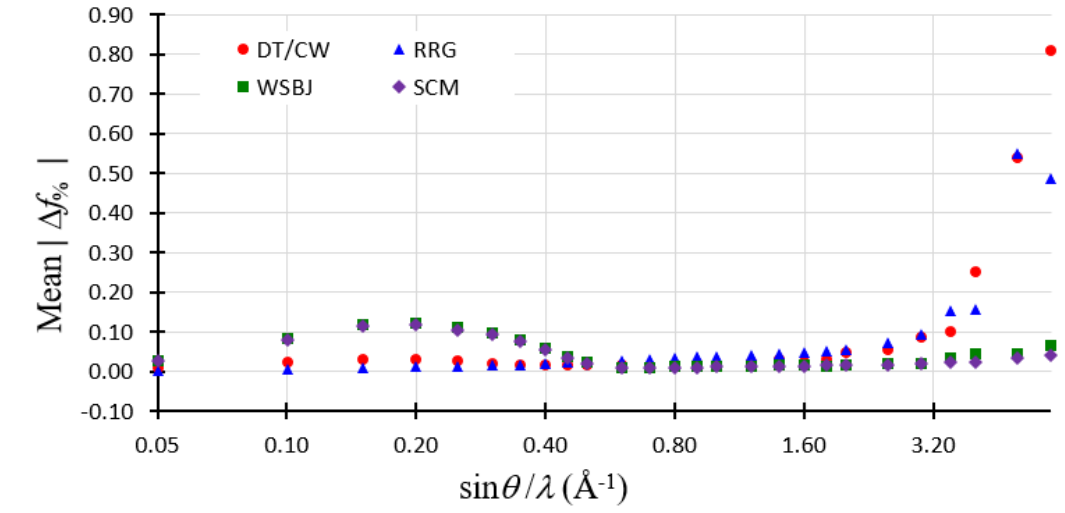


a)

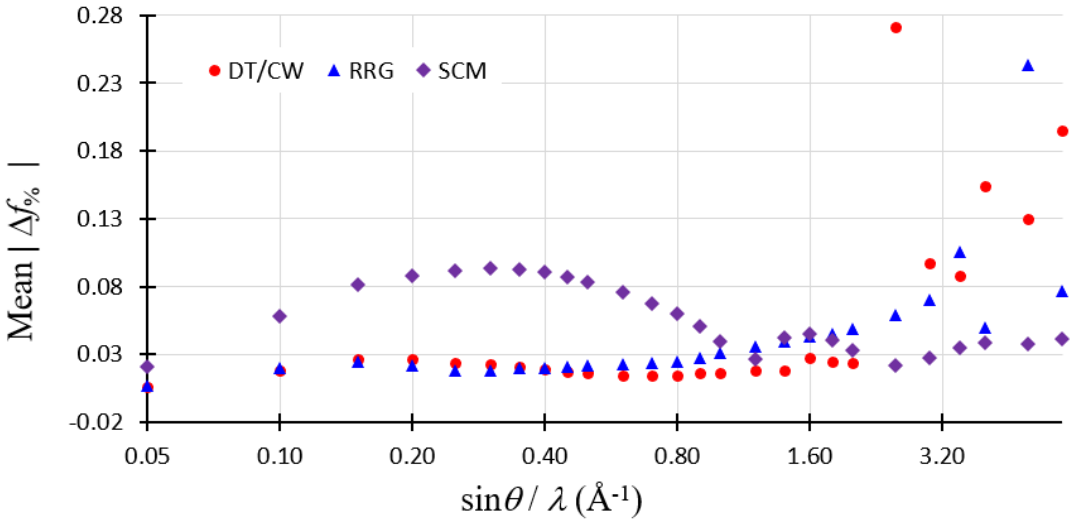


b)

Figure 8 . The *maximum percent* differences in the X-ray scattering factors ($\Delta f_{\%}$) between this work and the previous studies for each $\sin\theta/\lambda$ grid point between 0 and 6 Å^{-1} (Doyle & Turner, 1968) for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Rn (Z=86). The elements Tc (Z=43) and Pm (Z=61) were excluded as no data are given in the RRG study for these species. The previous studies are identified as follows: ● DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). For convenience, a logarithmic base-2 scale is used for the x-axis in both graphs. The symbols above each graph identify elements for which the maximum percent differences $\Delta f_{\%}$ were observed at the $\sin\theta/\lambda$ grid points.

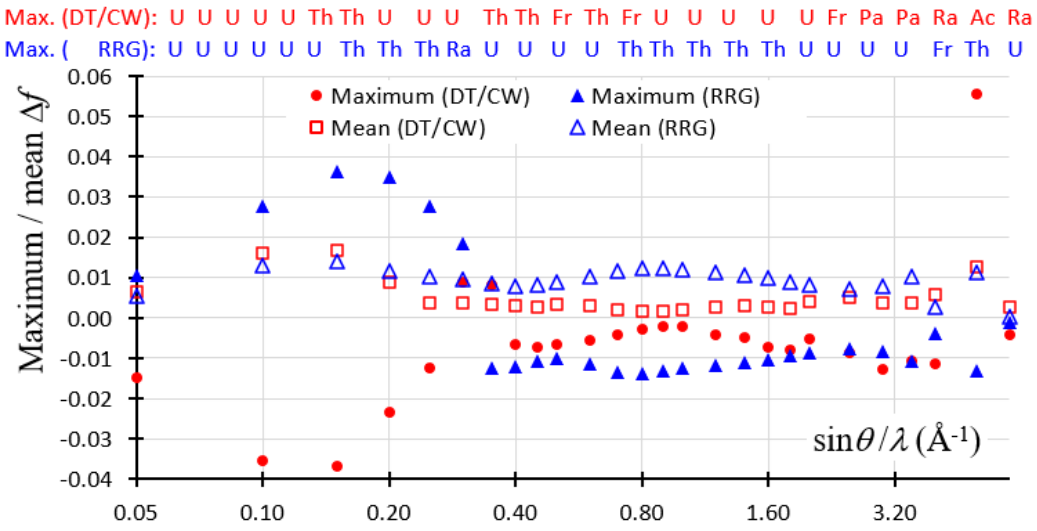


a)

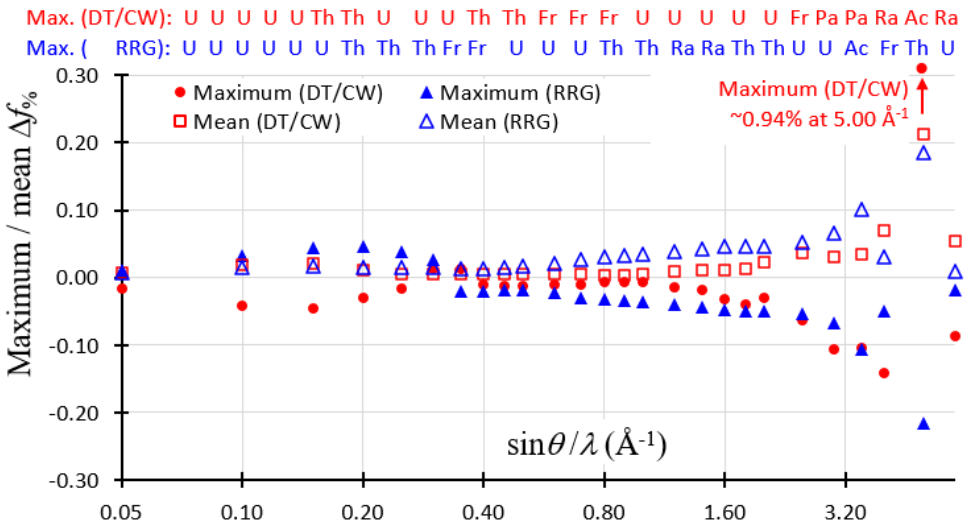


b)

Figure 9 . The *mean absolute percent* differences in the X-ray scattering factors ($|\Delta f_{\%}|$) between this work and the previous studies for **a)** B (Z=5) – Ar (Z=18) and **b)** K (Z=19) – Rn (Z=86) plotted as a function of $\sin\theta/\lambda$. The elements Tc (Z=43) and Pm (Z=61) were excluded as no data were given in the RRG study for these species. The averaging at each $\sin\theta/\lambda$ grid point between 0 and 6 Å^{-1} (Doyle & Turner, 1968) was done over a specified group of atoms: **a)** B – Ar, and **b)** K – Rn. The previous studies are identified as follows: ● DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968), ▲ RRG - Rez, Rez & Grant (1994), ■ WSBJ - Wang, Smith, Bunge & Jáiregui (1996), and ♦ SCM - Su & Coppens (1997) and Macchi & Coppens (2001). For convenience, a logarithmic base-2 scale is used for the *x*-axis in both graphs.



a)



b)

Figure 10 **a)** The *maximum* and *mean* differences (Δf) and **b)** the *maximum* and *mean percent* differences in the X-ray scattering factors between this work and the previous studies for each $\sin\theta/\lambda$ grid point between 0 and 6 Å^{-1} (Doyle & Turner, 1968) for Fr (Z=87) – U (Z=92). The mean quantities were obtained be averaging over all elements between Fr and U. The previous studies are identified as follows: DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968), and RRG - Rez, Rez & Grant (1994). The symbols above each graph identify elements for which the *maximum* **a)** differences Δf and **b)** percent difference $\Delta f_{\%}$ were observed at the $\sin\theta/\lambda$ grid points. For convenience, a logarithmic base-2 scale is used for the x-axis in both graphs.

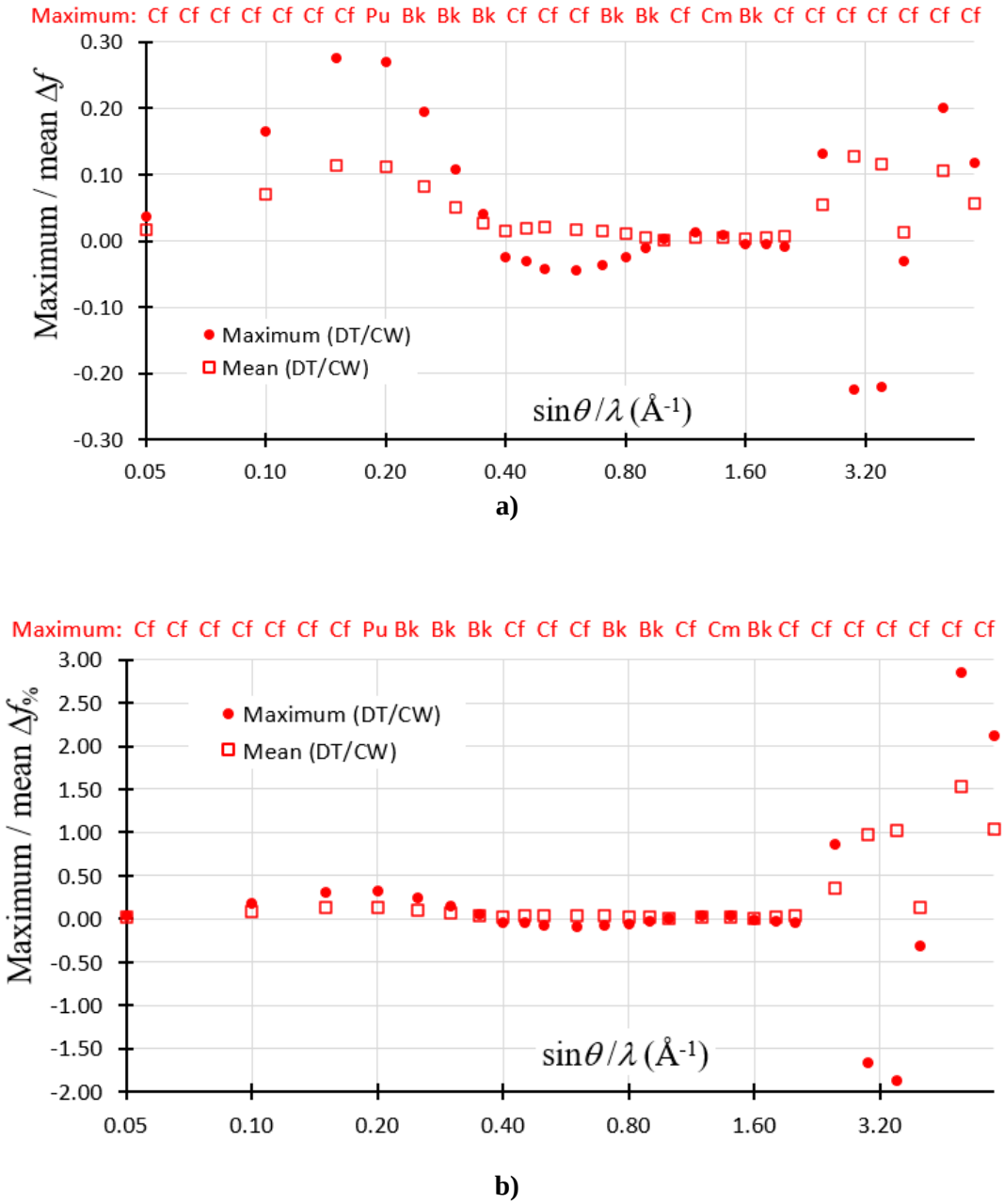
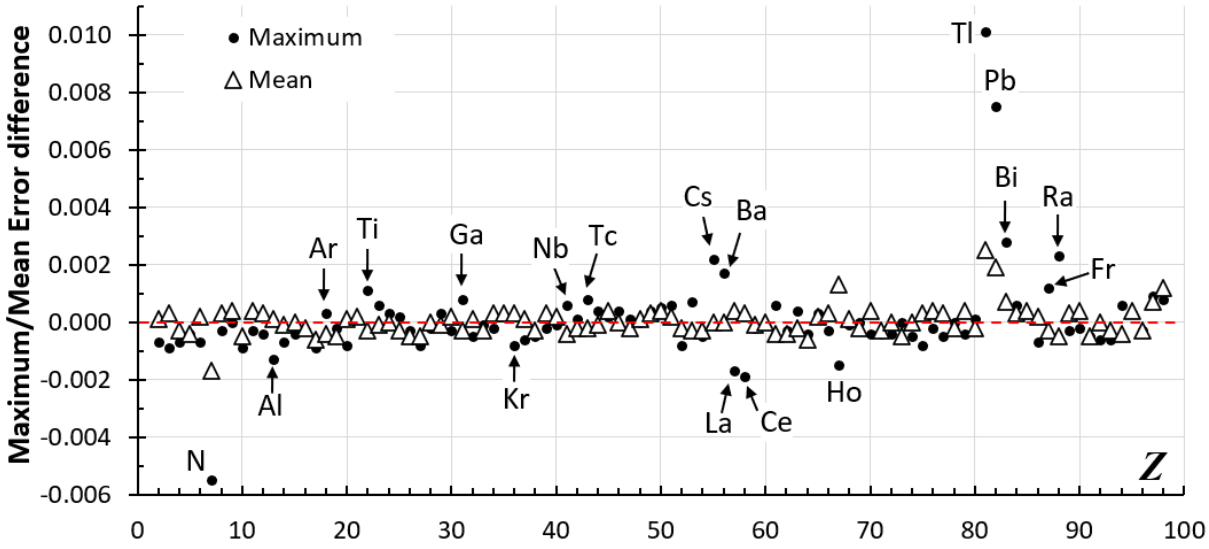
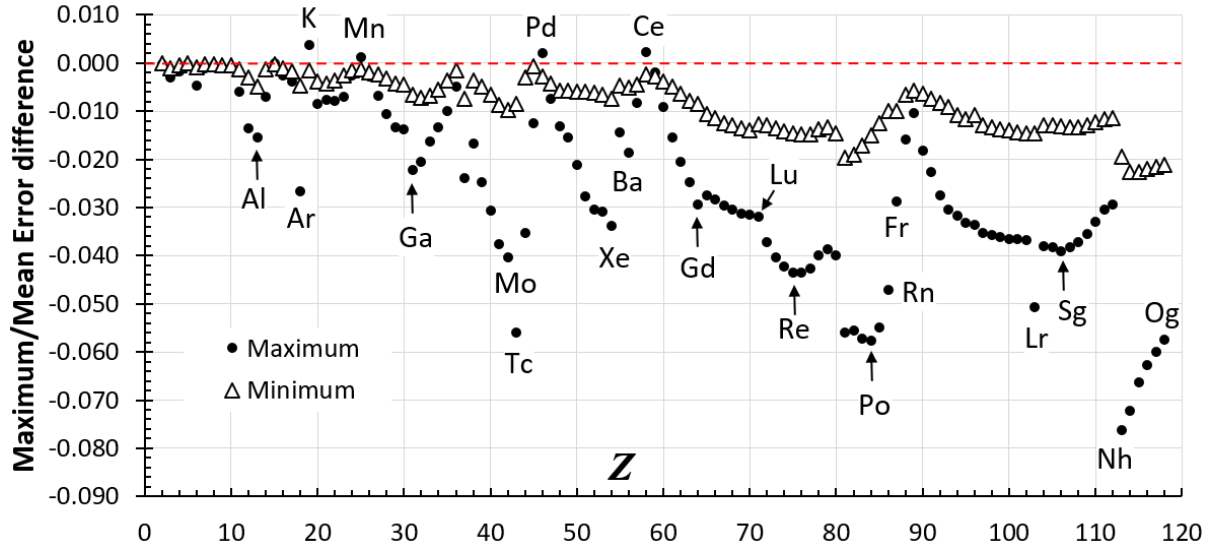


Figure 11 **a)** The *maximum* and *mean* differences (Δf) and **b)** the *maximum* and *mean percent* differences ($\Delta f_{\%}$) in the X-ray scattering factors between this work and the previous study for each $\sin\theta/\lambda$ grid point between 0 and 6 Å^{-1} (Doyle & Turner, 1968) for Np (Z=93) – Cf (Z=98). The mean quantities were obtained be averaging over all elements between Np and Cf. The previous study is identified as follows: DT/CW - Doyle & Turner (1968) and Cromer & Waber (1968). The symbols above each graph identify elements for each the *maximum* **a)** differences Δf and **b)** percent differences $\Delta f_{\%}$ were observed at the $\sin\theta/\lambda$ grid points. For convenience, a logarithmic base-2 scale is used for the x-axis in both graphs.



a)



b)

Figure 12 **(a)** The differences in the maximum and mean errors of the interpolating function (39) with $m = 4$ for atoms with Z = 2 – 98 between this work and the literature (Doyle & Turner, 1968; Cromer & Waber, 1968; Maslen, Fox & O'Keefe, 2006). Negative values indicate a lower error (better fit) in this work. **(b)** The differences in the maximum and mean errors of the interpolating function (39) with $m = 4$ and $m = 5$ (both from this study) for atoms with Z = 2 – 118. Negative values indicate a lower error in the $m = 5$ expansion. The interpolated $\sin\theta\,/\lambda$ range for both fits is $0 - 2$ Å^{-1}.

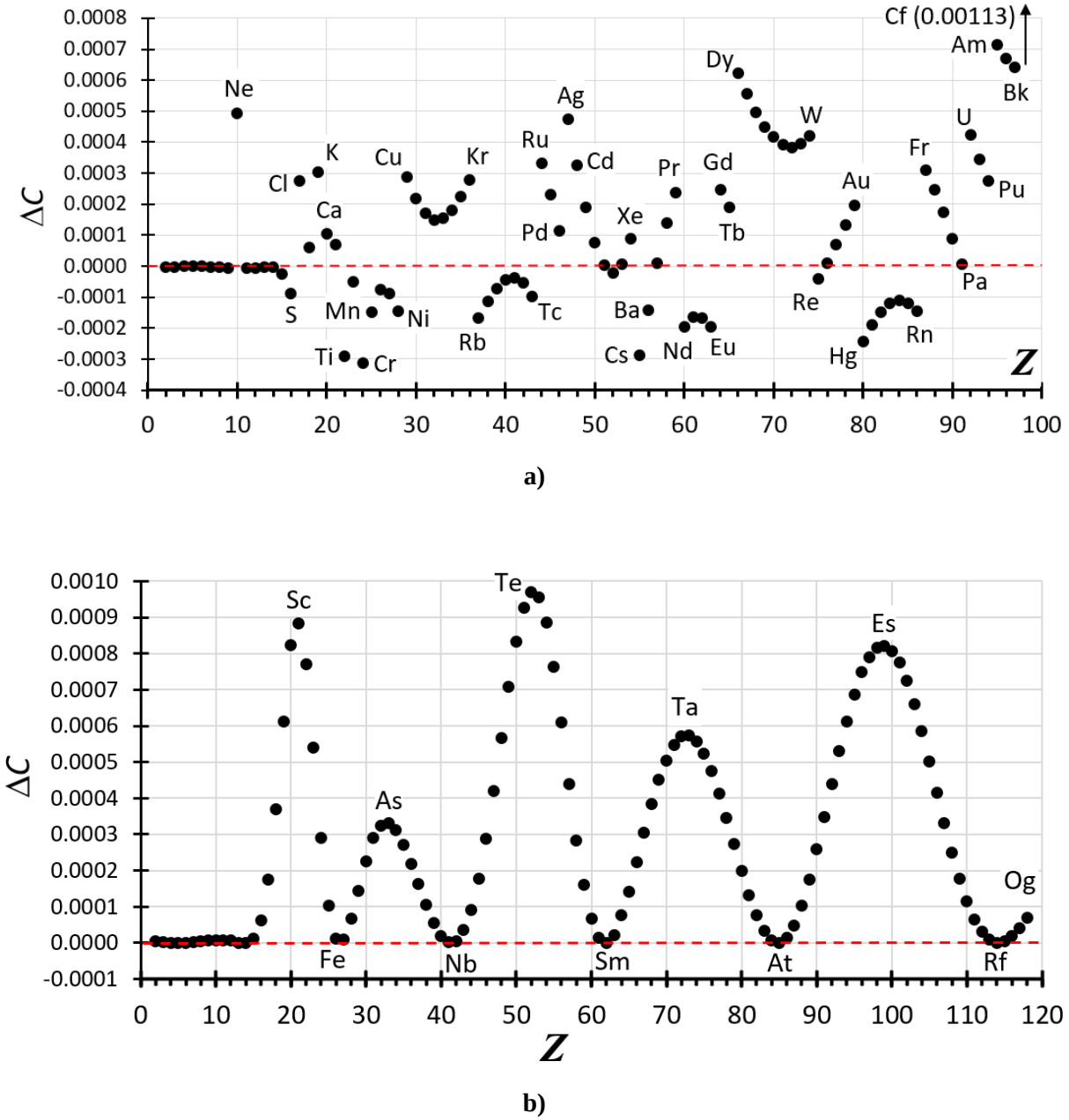
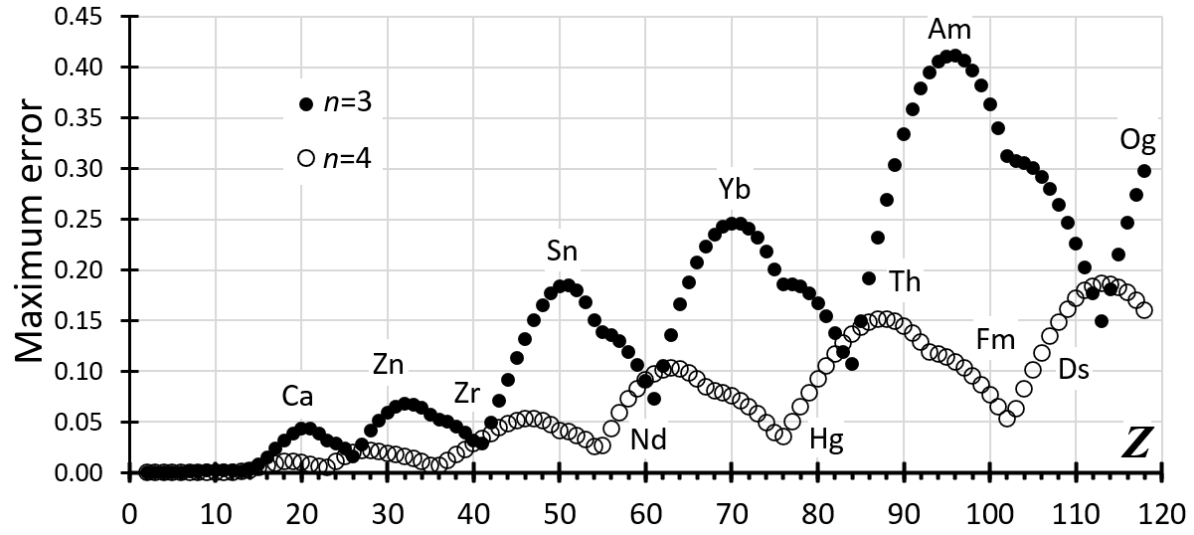
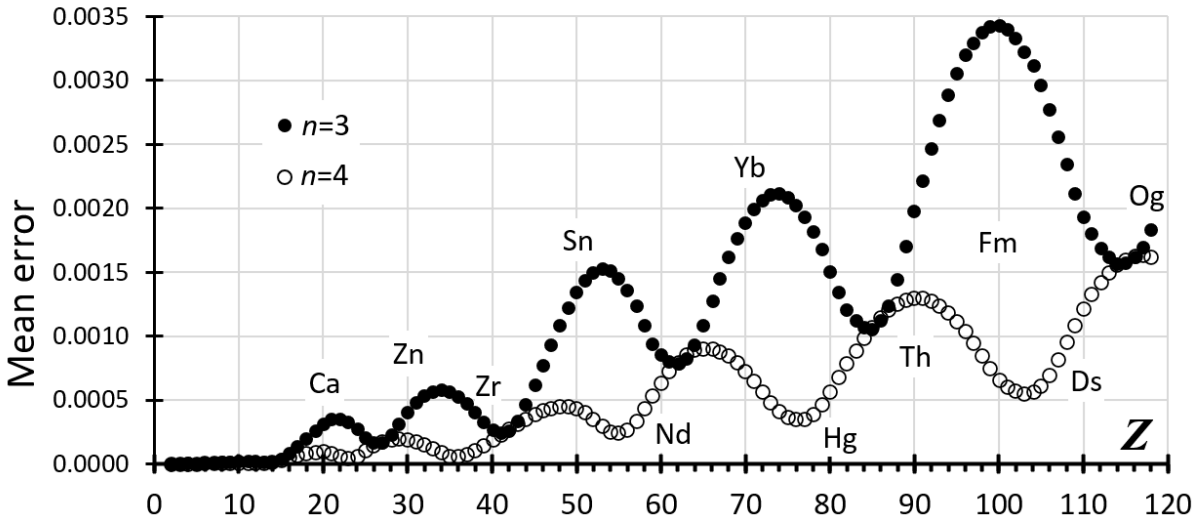


Figure 13 **(a)** The differences in the correlation coefficient C (ΔC) of the interpolating function (40) with $n = 3$ for atoms with Z = 2 – 98 between this work and the literature (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006). Positive values indicate higher C (better fit) in this work. **(b)** The differences in the correlation coefficient C (ΔC) of the interpolating function (40) with $n = 3$ and $n = 4$ (both from this study) for atoms with Z = 2 – 118. Positive values indicate a higher C (better fit) in the $n = 4$ expansion. The interpolated interval is $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} in both fits.



a)



b)

Figure 14 The **(a)** maximum and **(b)** mean errors of the interpolating function (40) with $n = 3$ (●) and $n = 4$ (○) for the atomic X-ray scattering factors in the $2 \leq \sin\theta\,/\lambda \leq 6$ Å^{-1} interval.

References

Autschbach, J. (2012). *J. Chem. Phys*. **136**, 150902-1 – 150902-15.

Bachau, H., Cormier, E., Decleva, P., Hansen, J. E. & Martín, F. (2001). *Rep. Prog. Phys.* **64**, 1815 - 1942.

Bevington, P. & Robinson, D. K. (2002). *Data Reduction and Error Analysis for the Physical Sciences*, 3rd edition. McGraw-Hill Science/Engineering/Math.

Byrd, R. H., Lu, P., Nocedal, J. & Zhu, C. (1995). *SIAM J. Sci. Comput.* **16**, 1190 – 1208.

Corana, A., Marchesi, M., Martini, C. & Ridella, S. (1987). *ACM Trans. Math. Softw*. **13**, 262 – 280.

Coulthard, M. A. (1967). *Proc. Phys. Soc.* **91**, 44 – 49.

Coppens, P. (1997). *X-ray Charge Densities and Chemical Bonding.* Oxford University Press, New York.

CRC Handbook of Chemistry and Physics (2021), 102nd edition, pages 1-1 – 1-15, Rumble Jr, J. R., Bruno, T. J. & Doa, M. J. editors, CRC Press.

Cromer, D. T. & Waber, J. T. (1968). Unpublished work reported in *International Tables for X-ray Crystallography* (1974), Vol. IV, p. 71. Birmingham: Kynoch Press. (Present distributor: Kluwer Academic Publishers, Dordrecht.)!

de Boor, C. (1977). *SIAM J. Numerical Analysis* **14**, 441 – 472.

Dennis, J. E., Gay, D. M. & Welsch, R. E. (1981a). *ACM Trans. Math. Softw*. **7**, 348 – 368.

Dennis, J. E., Gay, D. M. & Welsch, R. E. (1981b). *ACM Trans. Math. Softw*. **7**, 369 – 383.

Dennis, J. E. & Schnabel, R. B. (1996). *Numerical Methods for Unconstrained Optimization Nonlinear Equations*, SIAM, Philadelphia, PA.

Desclaux, J. P. (1973). *At. Nucl. Data. Tabl.* **12**, 311 – 406.

Desclaux, J. P. (1975). *Comput. Phys. Comm.* **9**, 31 – 45.

Desclaux, J. P. & Fricke, B. (1980). *J. Physique.* **41**. 943 – 946.

Desclaux, J. P., Mayers, D. F. & O'Brien, F. (1971). *J. Phys. B. Atom. Molec. Phys*. **4**, 631 – 642.

Doyle, P. A. & Turner, P. S. (1967). *Acta Cryst.* **22**, 153.

Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* A**24**, 390 – 397.

Dyall K. G., Grant, I. P., Johnson, C., Parpia, F.A. & Plummer, E. (1989). *Comput. Phys. Comm*. **55**, 425 – 456.

Fletcher, R. (1987). *Practical Methods of Optimization*, 2nd edition, John Wiley & Sons.

Fong, K. W., Jefferson, T. H., Suyehiro, T. & Walton, L. (1993). *Guide to the SLATEC Common Mathematical Library*. https://www.netlib.org/slatec/guide

Fox, A. G., O'Keefe, M. A. & Tabbernor, M. A. (1989). *Acta Cryst.* A**45**, 786 – 793.

Fox, P. A. (1984). *The PORT Mathematical Subroutine Library*. AT&T Bell Telephone Laboratories, Inc. http://www.netlib.org/port/

Froese Fischer, C. (2007). *Adv. At. Mol. Opt. Phys*. **55**, 235 – 291.

Froese Fischer, C. (2021). *Atoms.* **9**, 50-1 – 50-14.

Froese Fischer, C., Gaigalas, G., Jönsson, P. & Bieroń, J. (2019). *Comput. Phys. Comm.* **237**, 184 – 187. https://www-amdis.iaea.org/GRASP2K/

Gay, D. M. (1981). *SIAM J. Sci. Statist. Comput.* **2**, 186 – 197.

Gay, D. M. (1983). *ACM Trans. Math. Softw.* **9**, 139.

Giacovazzo, C., Monaco, H. L., Viterbo, D., Scordari, F., Gilli, G., Zanotti, G. & Catti, M. (1992). *Fundamentals of Crystallography*. Oxford University Press, New York.

Gilbert, J. C. & Nocedal, J. (1992). *SIAM J. Optim.* **2**, 21 – 42.

Grant, I. P. (1961). *Proc. Roy. Soc. A* **262**, 555 – 576.

Grant, I. P. (1970). *Adv. Phys.* **19**, 747 – 811.

Grant, I. P., Mayers, D. F. & Pyper, N. C. (1976). *J. Phys B: At. Mol. Phys.* **9**, 2777 – 2796.

Grant, I. P., McKenzle, B. J., Norrington, P. H., Mayers, D. F. & Pyper, N. C. (1980). *Comp. Phys. Comm.* **21**, 207 – 231.

Gu, M. F. (2008). *Can. J. Phys.* **86**, 675 – 689. https://www-amdis.iaea.org/FAC/

Guerra, M., Amaroa, P., Santos, J. P. & Indelicato, P. (2017). *At. Nucl. Data. Tabl.* **117-118**, 439 – 457.

Indelicato, P. & Desclaux, J. P. (1990). *Phys. Rev. A* **42** 5139 – 5149.

Johnson, W. R. & Soff, G. (1985). *At. Nucl. Data. Tabl.* **33**, 405 – 446.

Jones, R. E. (1981). *Subroutine GAUS8 in SLATEC Common Mathematical Library*, Version 4.1, July 1993. https://www.netlib.org/slatec/

Jönsson, P., He, X., Froese Fischer, C. & Grant, I. P. (2007). *Comput. Phys. Comm.* **177**, 597 – 622.

Jönsson, P., Gaigalas, G., Bieroń, J., Froese Fischer, C. & Grant, I. P. (2013). *Comput. Phys. Comm.* **184**, 2197 – 2203.

Kahl, E. V. & Berengut, J. C. (2019). *Comput. Phys. Comm.* **238**, 232 – 243. https://github.com/drjuls/AMBiT

Levenberg, K. (1944). *Quart. Appl Math.* **2**, 164 – 168.

Macchi, P. & Coppens, P. (2001). *Acta Cryst.* A**57**, 656 – 662.

Mann, J. B. (1968). Unpublished work reported in *International Tables for X-ray Crystallography* (1974), Vol. IV, p. 71. Birmingham: Kynoch Press. (Present distributor: Kluwer Academic Publishers, Dordrecht.)

Marquardt, D.W. (1963). *J. Soc. Indust. Appl. Math.* **11**, 431 – 441.

Martin, W. C. (2021). in *CRC Handbook of Chemistry and Physics*, p. 1–16, 102nd edition, Rumble Jr, J.R., Bruno, T. J. and Doa, M. J. editors, CRC Press.

Maslen, E. N., Fox, A. G. & O'Keefe, M. A. (2006). In *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589.

Matsuoka, O. & Watanabe, Y. (2001). *Comput. Phys. Comm.* **139**, 218 – 234.

Morales, J. L. & Nocedal, J. (2011). *ACM Trans. Math. Softw.* **38**, 1 – 4.

Moré, J., Garbow, B. & Hillstrom, K. (1981). *ACM Trans. Math. Softw.* **7**, 136 – 140.

Moré, J., Garbow, B. & Hillstrom, K. (1999). *MINPACK - software for solving nonlinear equations and nonlinear least squares problems*. University of Chicago, Argonne National Laboratory, USA. https://www.netlib.org/minpack/

Moré, J. J., Sorensen, D. C., Hillstrom, K. E. & Garbow, B. S. (1984). *The MINPACK Project*, in *Sources and Development of Mathematical Software*, ed. Cowell, W. J., Prentice-Hall, pp. 88 – 111.

Parpia, E A., Froese Fischer, C. & Grant, I. P. (1996). *Comput. Phys. Comm*. **94**, 249 – 271.

Piessens, R. & de Doncker, E. (1980). *Subroutine QAG in SLATEC Common Mathematical Library*, Version 4.1, July 1993. https://www.netlib.org/slatec/

Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3rd edition. Cambridge University Press.

Pyper, N. C. (2020). *Phil. Trans. R. Soc. A* **378**, 20190305.

Pyykkö, P. (2012). *Annu. Rev. Phys. Chem.* **63**, 45 – 64.

Rafferty, J. & Norling, R. (1986). *Cricket Graph Version 1.0*. Cricket Software Inc., Philadelphia, PA 19104, USA.

Rez, D., Rez, P. & Grant, I. (1994). *Acta Cryst.* A**50**, 481 – 497.

Rez, D., Rez, P. & Grant, I. (1997). *Acta Cryst.* A**53**, 522.

Schnabel, R., Koontz, J. & Weiss, B. (1982). *A modular system of algorithms for unconstrained minimization*, Technical Report CU-CS-240-82, Computer Science Department, University of Colorado at Boulder, Boulder, CO.

Schwarzenbach, D. (1996). *Crystallography*. Wiley, Chichester.

Su, Z. & Coppens, P. (1997). *Acta Cryst.* A**53**, 749 – 762.

Su, Z. & Coppens, P. (1998). *Acta Cryst.* A**54**, 357.

Swirles, B. (1935). *Proc. Roy. Soc. A* **152**, 625 – 649.

Tatewaki, H., Yamamoto, S. & Hatano, Y. (2017). *ACS Omega* **2**, 6072 – 6080.

Vandevender, W. H. & Haskell, K. H. (1982). *SIGNUM Newsletter*, **17**, 16 – 21.

Visscher, L. & Dyall, K. G. (1997). *At. Nucl. Data. Tabl.* **67**, 207 – 224.

WebElements: The periodic table on the WWW. http://www.webelements.com Accessed 15 June 2022.

Zatsarinny, O. & Froese Fischer, C. (2016). *Comput. Phys. Comm.* **202**, 287 – 303. https://github.com/compas/dbsr_hf

Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. (1994). *Tech. Report, NAM-11, EECS* Department, Northwestern University, 1994.

Wang, J., Smith Jr,V. H., Bunge, C. F. & Jáiregui, R. (1996). *Acta Cryst*. A**52**, 649 – 658.

Weissbluth, M. (1978). *Atoms and Molecules*. Academic Press, Inc., New York / London.

Wolfram Research, Inc., (2022) *Mathematica, Version 13.0*, Champaign, IL, USA.

Chapter 3. ARTICLE 2 - Revisited Relativistic Dirac-Hartree-Fock X-ray Scattering Factors. II. Chemically Relevant Cations and Selected Monovalent Anions for Z = 2-112.

Authors

Shiroye Olukayode[a]*, Charlotte Froese Fischer[b] and Anatoliy Volkov[a]*

[a]Department of Chemistry and Computational Science Program, Middle Tennessee State University, Murfreesboro, TN, 37132, USA
[b]Department of Computer Science, University of British Columbia, 2366 Main Mall, Vancouver, British Columbia, V6T1Z4, Canada

Correspondence email: oos2b@mtmail.mtsu.edu; anatoliy.volkov@mtsu.edu

Synopsis Fully relativistic X-ray scattering factors for all chemically relevant cations, the monovalent anions of the halogens, the excited (valence) ns^1np^3 states of carbon and silicon, and several exotic cations for atoms with $Z > 104$ have been determined in the $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range using the B-spline Dirac-Hartree-Fock method of Zatsarinny & Froese Fischer (2016) [*Comput. Phys. Comm.* **202**, 287 – 303]. The study also reports the analytical conventional and extended interpolating functions for the 0 – 2 and 2 – 6 Å^{-1} $\sin\theta/\lambda$ intervals, and includes a thorough comparison with the results from the earlier investigations.

Abstract The previously described approach for determination of the atomic X-ray scattering factors at the Dirac-Hartree-Fock level [Olukayode, Froese Fischer & Volkov, A. (2022) *Acta Cryst.* A, submitted] has been used to evaluate the X-ray scattering factors for a total of 318 species including all chemically-relevant cations [Greenwood & Earnshaw (1997). *Chemistry of the Elements*], six monovalent anions (O^-, F^-, Cl^-, Br^-, I^-, At^-), the ns^1np^3 excited (valence) states of carbon and silicon, and several exotic cations (Db^{5+}, Sg^{6+}, Bh^{7+}, Hs^{8+} and Cn^{2+}) for which the chemical compounds have been recently identified, thus significantly extending the coverage of species relative to all the earlier studies. Unlike the currently recommended IUCr data [Maslen, Fox & O'Keefe (2006). In *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589] which originate from different levels of theory including the non-relativistic Hartree-Fock and correlated methods, as well as the relativistic Dirac-Slater calculations, the re-determined X-ray scattering factors come from a uniform treatment of all species within the same relativistic B-spline Dirac-Hartree-Fock approach [Zatsarinny & Froese Fischer (2016). *Comput. Phys. Comm.* **202**, 287 – 303] that includes the Breit interaction correction and the Fermi nuclear charge density model. While it was not possible to compare the quality of the generated wavefunctions with those from the previous studies due to a lack (to the best of our knowledge) of such data in the literature, a careful comparison of the total electronic energies and the estimated ionization energies with the experimental and theoretical values from other studies instils confidence in the quality of the calculations. A combination of the B-spline approach and a fine radial grid allowed for a precise determination of the X-ray scattering factors for each species in the entire $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range thus avoiding the necessity for extrapolation in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} which, as was shown in our first study, may lead to inconsistencies. In contrast to the Rez, Rez & Grant [(1994). *Acta Cryst.* A**50**, 481 – 497] work, no additional approximations were introduced when calculating the wavefunctions of the anions. The conventional and extended expansions were employed to

produce interpolating functions for each species in both the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} intervals, with the extended expansions offering a significantly better accuracy at a minimal computational overhead. The outcomes of the undertaken research should be of interest to the members of crystallographic community who push the boundaries of the accuracy and precision of the X-ray diffraction studies.

Keywords: Relativistic X-ray scattering factors for ions; relativistic Dirac-Hartree-Fock; interpolation of X-ray scattering factors.

3.1 Introduction

In the previous paper (Olukayode, Froese Fischer & Volkov, 2022; called Paper I in the following) we presented a straightforward and what appears to be a reliable procedure for determination of the atomic X-ray scattering factors at the relativistic Dirac-Hartree-Fock level (Swirles, 1935; Grant, 1961, 1970) using the B-spline DBSR_HF code recently developed by Zatsarinny & Froese Fischer (2016). The DBSR_HF-based wavefunctions for the ground-state electronic configurations of all elements with the atomic number Z = 2 (He) – 118 (Og) were determined using the extended average level (EAL) approximation (Grant, Mayers & Pyper, 1976; Grant et al., 1980; Dyall et al., 1989), and included both the Breit interaction correction and the Fermi nuclear charge density model (Johnson & Soff, 1985; Zatsarinny & Froese Fischer, 2016).). The X-ray scattering factors for each species were evaluated in the entire $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range, and were interpolated using the conventional (Maslen, Fox & O'Keefe, 2006) and extended expansions in the separate $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} intervals using the newly developed Fortran code SF (Paper I). In this study the approach is extended to the treatment of cations and the selected monovalent anions.

To the best of our knowledge, the first more or less complete set of the relativistic X-ray scattering factors for ions were calculated by Cromer & Waber (1965) using the Dirac-Slater approach (Liberman, Waber & Cromer, 1965). The calculations included 101 neutral atoms (Li – No) and 107 ions ranging from H^- to Am^{3+} including five anions (H^-, F^-, Cl^-, Br^-, and I^-). The scattering factors for each species were calculated to two decimal digits on a fairly coarse $\sin\theta/\lambda$ grid between 0 and 2 Å^{-1}. Because at this time those values are considered essentially obsolete, we shall exclude the results of the Cromer & Waber (1965) study from further discussions.

The earliest set of the relativistic Dirac-Hartree-Fock scattering factors that are still being used today (Maslen, Fox & O'Keefe, 2006) were determined by Doyle & Turner (1968)

for $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} to three decimal digits using wavefunctions obtained from the Coulthard (1967) code; these are labelled as "RHF" in Table 6.1.1.3 in volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). Unfortunately, the Doyle & Turner (1968) calculations neglected the Breit interaction correction and the finite radius of the atomic nucleus, and due to convergence issues, only the selected ions (Li^+, Be^{2+}, Na^+, Mg^{2+}, Cl^-, K^+, Ca^{2+}, V^{2+}, Mn^{2+}, Fe^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^+, Zn^{2+}, Br^-, Rb^+, Sr^{2+}, Sn^{2+}, Sn^{4+}, I^-, Cs^+) were successfully processed.

The IUCr-recommended X-ray scattering factors for H^-, Li^+ and Be^{2+} come from non-relativistic but correlated wavefunctions of Thakkar & Smith (1992) and labelled as "C" in Maslen, Fox & O'Keefe (2006). The values of Thakkar & Smith (1992) for Li^+ and Be^{2+} are essentially identical to those of Doyle & Turner (1968) but because the two sets of values are reported to a different number of decimal places (four and three, respectively), the mean average difference over the common $\sin\theta/\lambda$ grid points is 0.0004 atomic units for both ions. It suggests that the relativistic and correlation effects are not important for these species due to a low atomic weight and the presence of only two electrons, respectively. In addition, the Thakkar & Smith (1992) X-ray scattering factors for H^-, Li^+ and Be^{2+} were calculated for $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} using a finer grid (especially in the low $\sin\theta/\lambda$ region) than that used by Doyle & Turner (1968).

For the remaining ions between O^- and Ge^{4+} and the ns^1np^3 excited (valence) states of carbon and silicon, denoted in Maslen, Fox & O'Keefe (2006) as C_{val} and Si_{val}, respectively, the X-ray scattering factors were calculated by Cromer & Mann (1968a) using the non-relativistic numerical Hartree-Fock wavefunctions of Mann (1967), and are labelled as "HF" in Maslen, Fox & O'Keefe (2006). These were evaluated to three decimal digits on a fine $\sin\theta/\lambda$ grid between 0 and 1.5 Å^{-1}. For the higher $\sin\theta/\lambda$ values, the relativistic Dirac-Hartree-Fock scattering factors for neutral atoms of Doyle & Turner (1968) or Cromer &

Waber (1968) are supposed to be used. The latter study accounted for a finite size of the nucleus using the uniform charge distribution model. However, while the Doyle & Turner (1968) X-ray scattering factors cover the entire 0 – 6 Å^{-1} range, the Cromer & Waber (1968) data terminate at 2 Å^{-1} and were extrapolated to 6 Å^{-1} by Fox, O'Keefe & Tabbernor (1989). Unfortunately, as demonstrated in Paper I, the extrapolation leads to some numerical inconsistencies.

Finally, for all the remaining ions (Maslen, Fox & O'Keefe, 2006) the X-ray scattering factors were evaluated by Cromer & Waber (1968) using the relativistic Dirac-Slater method for the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range, and are labelled as "*DS" in Maslen, Fox & O'Keefe (2006). Beyond 2 Å^{-1}, the X-ray scattering factors for ions need to be augmented by the relativistic Dirac-Hartree-Fock values for neutral atoms from the Doyle & Turner (1968) (RHF) study and by the extrapolated values of Fox, O'Keefe & Tabbernor (1989). The X-ray scattering factors of both Cromer & Waber (1968) and Fox, O'Keefe & Tabbernor (1989) were reported to three decimal digits.

According to *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006), the X-ray scattering factors of ions, C_{val}, and Si_{val} have been interpolated in $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range (Doyle & Turner, 1968; Cromer & Mann, 1968b; Cromer & Waber, 1968; Thakkar & Smith, 1992) using the expression

$$f(\sin\theta/\lambda) = \sum_{i=1}^{m} a_i \exp(-b_i \sin\theta^2/\lambda^2) + c \tag{50}$$

with $m = 4$ as proposed by Vand, Eiland & Pepinsky (1957). For $\sin\theta/\lambda \geq 2$ Å^{-1}, the interpolating functions of Fox, O'Keefe & Tabbernor (1989)

$$f(\sin\theta/\lambda) = \exp\left(\sum_{i=0}^{n} a_i (\sin\theta/\lambda)^i\right) \tag{51}$$

with $n = 3$ as determined for neutral atoms are employed.

The complexity of the existing X-ray scattering data for ions was recognized previously by Rez, Rez & Grant (1994), Wang et al.(1996), and Macchi & Coppens (2001).

Rez, Rez & Grant (1994) used the multiconfiguration relativistic Dirac-Hartree-Fock to rectify the problem. Their calculations were performed in the MCDF program (Grant et al., 1980), which later became known as GRASP (Dyall et al., 1989), using the extended average level (EAL) model (Grant, Mayers & Pyper, 1976; Grant et al., 1980; Dyall et al., 1989). The resulting X-ray scattering factors were calculated on the same $\sin\theta/\lambda$ grid as used in the Doyle & Tuner (1968) study and were reported to four decimal places. Regrettably, it is not known which nuclear model was used by Rez, Rez & Grant (1994) and whether the Breit interaction correction was applied. In addition, their work covered only a subset of ions calculated in the earlier studies (Cromer & Waber, 1968) though they did include the X-ray scattering factors for Cr^{4+} and O^{2-} that had not been reported before. The calculations of Dirac-Hartree-Fock wavefunctions for anions (O^{2-}, F^{-}, Cl^{-}, Br^{-}, and I^{-}) in the Rez, Rez & Grant (1994) work required the use of the Watson sphere approximation (Watson, 1958) which involves surrounding of anions by a sphere of positive charge. Finally, Rez, Rez & Grant (1994) used expansion (50) with $m = 4$ with $c = 0$ to produce both a conventional fit in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range and a "*lower-accuracy fit over an extended range to 6.0 Å^{-1}*".

Wang et al. (1996) also used the multiconfiguration relativistic Dirac-Hartree-Fock method but their calculations included only eight species: six ions (Li^{+}, Be^{2+}, F^{-}, Na^{+}, Mg^{2+} and Cl^{-}) and the excited (valence) states of carbon (C_{val}) and silicon (Si_{val}), though no additional approximations were included when calculating the two anions. The X-ray scattering were computed to four decimal digits on a fine $\sin\theta/\lambda$ grid between 0 and 6 Å^{-1}, but no interpolating expansions were provided.

Macchi & Coppens (2001) extended the work of Su & Coppens (1997; 1998a) for neutral atoms to most "*chemically important ions up to I*$^{-}$" except for Mo^{5+} for which

"*convergence could not be achieved*". These studies used the multiconfiguration relativistic Dirac-Hartree-Fock method in the optimal level (OL) mode as implemented in the program GRASP92 (Parpia et al., 1996) and included the Fermi nuclear charge model, though it is not clear from the publication whether the Breit interaction correction was applied to the generated wavefunctions. The X-ray scattering factors were produced for the $0 \leq \sin\theta/\lambda \leq 10$ Å^{-1} range with a grid spacing of 0.05 Å^{-1} and were reported to nine decimal digits (Macchi & Coppens, 2001). While the original Su & Coppens (1997) work mentions that the X-ray scattering factor integrals were evaluated numerically to five decimal digits using a combination of the cubic spline interpolation and the composite-Simpson approach (Burden & Faires, 1989), a careful inspection of the $f(0)$ values in the Macchi & Coppens (2001) data suggests that at least in the low $\sin\theta/\lambda$ region, the precision of the numerically integrated X-ray scattering factors is indeed 8 – 9 decimal digits. However, from our experience (Paper I), precision of the numerical integration usually deteriorates with increasing $\sin\theta/\lambda$. As in the work of Wang et al. (1996), no additional approximations were introduced when calculating the monovalent anions (O^-, F^-, Cl^-, Br^-, and I^-) but unlike the Rez, Rez & Grant (1994) study, the O^{2-} and Cr^{4+} ions were not calculated. Finally, in contrast to the adopted conventions (Maslen, Fox & O'Keefe, 2006), the interpolating functions for all the calculated ions included expansion (50) with $m = 6$ and the coefficient c set to zero for the three separate $\sin\theta/\lambda$ intervals of 0 – 2 Å^{-1}, 2 – 4 Å^{-1}, and 4 – 6 Å^{-1} as was proposed by Su & Coppens (1997; 1998a).

The overarching goal of the present work is to complete a unform treatment of the X-ray scattering factors at the relativistic Dirac-Hartree-Fock level for all chemically-relevant cations (Greenwood & Earnshaw, 1997) and the selected monovalent anions (O^-, F^-, Cl^-, Br^-, I^-, which were included in most of the earlier studies, plus At^-), thus significantly extending the list of ions that were treated in all the previous studies. The objectives of this study are to

d) integrate the X-ray scattering factors with a high precision at all points of a fine $\sin\theta\,/\lambda$ grid between 0 and 6 $Å^{-1}$ (Wang et al., 1996);

e) determine the conventional (Maslen, Fox & O'Keefe, 2006) interpolating functions (50) and (51) for $0 \leq \sin\theta\,/\lambda \leq 2\ Å^{-1}$ and $2 \leq \sin\theta\,/\lambda \leq 6\ Å^{-1}$ ranges, respectively, which will allow the users [should there be interest] include the expansions into X-ray diffraction software with only minor modifications;

f) optimize the extended interpolating functions (50) and (51) ($m = 5$ and $n = 4$, respectively) in order to increase the accuracy of the interpolated X-ray scattering factors.

3.2 Methods

3.2.1 Relativistic calculations

In this work, we have performed calculations of the X-ray scattering factors for the following 318 species:

a) all cations of the elements with Z = 3 (Li) – 104 (Rf) listed in Greenwood & Earnshaw (1997) (Figure 2.5, page 28) thus significantly extending the coverage as compared to Maslen, Fox & O'Keefe (2006), Rez, Rez & Grant (1994), and Macchi & Coppens (2001),

b) six monovalent anions: O^{-}, F^{-}, Cl^{-}, Br^{-}, I^{-}, At^{-};

c) excited (valence) states of carbon ([He]$2s^1 2p^3$) and silicon ([Ne]$3s^1 3p^3$) – these are denoted as C_{val} and Si_{val} in Cromer & Waber (1968), Wang et al. (1996), and Macchi & Coppens (2001),

d) several exotic cations beyond Rf (Z=104) for which chemical compounds are known: Db^{5+} (Z=105) (Gäggeler & Türler, 2014), Sg^{6+} (Z=106) (Hübener et al., 2001; Pershina, Kratz & Fricke, 2000; Gäggeler & Türler, 2014), Bh^{7+} (Z=107) (Eichler et al., 2000; Pershina, Kratz, Fricke & Bastug, 2000; Gäggeler & Türler, 2014), Hs^{8+} (Z=108) (Gäggeler & Türler, 2014) and Cn^{2+} (Z=112) (Chiera et al., 2015).

The ground non-relativistic electronic *LS* states of all the calculated species are listed in Table S1. The values for the cations were taken from the *NIST Atomic Spectra Database* (Kramida et al., 2021) and Rodrigues et al. (2004). A careful examination of the configurations reveals several discrepancies between the two sources even though the *NIST Atomic Spectra Database* (NASD) does include selected data from Rodrigues et. al. (2004):

1) For Os^{2+} (Z=76), the ground state electronic configuration listed in NSAD is [Xe] $4f^{14}$ $5d^6$ while Rodrigues et. al. (2004) report it as [Xe] $4f^{14}$ $5d^5$ $6s^1$. We used the

NSAD configuration as it comes from a recent experimental study (Azarov, Tchang-Brillet & Gayasov, 2018).

2) The ground state electronic configuration of Th^{2+} (Z=90) is given in NSAD as [Rn] $6d^1\ 5f^1$ and as [Rn] $6d^2$ in Rodrigues et. al. (2004). We used the NSAD value since it is based on a more recent experimental investigation (Redman, Nave & Sansonetti, 2014).

3) For Np^{7+} (Z=93), Pu^{6+} and Pu^{7+} (Z=94), we chose to use the configurations reported by Rodrigues et. al. (2004): [Xe] $4f^{14}\ 5d^{10}\ 6s^2\ 6p^6$, [Rn] $5f^2$ and [Rn] $5f^1$, respectively, as the NSAD data ([Xe] $4f^{14}\ 5d^{10}\ 6s^2\ 6p^5\ 5f^1$, [Xe] $4f^{14}\ 5d^{10}\ 6s^2\ 6p^5\ 5f^3$, and [Xe] $4f^{14}\ 5d^{10}\ 6s^2\ 6p^4\ 5f^3$) were taken from a fairly old theoretical study.

The electronic configuration of Cn^{2+}, [Rn] $5f^{14}\ 6d^{10}$, was modelled after that of Hg^{2+}, [Xe] $4f^{14}\ 5d^{10}$. The electronic configurations of the anions (O^- , F^-, Cl^-, Br^-, I^-, At^-) were formed by adding an extra electron to the partially occupied *p*-orbital. The electronic configurations for the excited (valence) states of carbon and silicon are specified above.

The relativistic Dirac-Hartree-Fock calculations were performed in the DBSR_HF program of Zatsarinny & Froese Fischer (2016) following the same procedure and using the same parameters as in the previous study (Olukayode, Froese Fischer & Volkov, 2022). As before, the energy and orbital convergence tolerance criteria were set to their default values of 10^{-10} and 10^{-7} atomic units, respectively. For Cu^+ and Yb^{2+} cations, the orbital convergence criterion had to be increased to 3×10^{-7}. Unfortunately, the original DBSR_HF code has a bug when working with anions. A temporary workaround involved calculating the neutral atom first and then using the neutral atom's radial functions as initial guess when calculating anion. Using this technique, we were able to successfully complete the calculations for not only the six anions listed above, but also for C^-, C^{4-}, Si^-, P^-, S^-, Ge^-, As^-, Se^-, Mo^-, Sn^-, Sb^-, Ir^-, Au^-,

and Po^{-}. However, since DBSR_HF was not able to calculate the more common multivalent anions such as N^{3-}, O^{2-}, Si^{4-}, P^{3-}, S^{2-}, Ge^{4-}, As^{3-} etc. (Greenwood & Earnshaw, 1997) we shall not discuss any of these species in this manuscript. The remaining monovalent and multivalent anions will be a subject of a separate investigation.

As in the previous study (Paper I), the DBSR_HF calculations were performed at the extended average level (Grant, Mayers & Pyper, 1976; Grant et al., 1980; Dyall et al., 1989), and included the Breit correction and the nuclear charge density Fermi distribution function, but the quantum electrodynamical self-energy and vacuum polarization were applied to the energies only. The extended average level model (EAL), which is essentially identical to the average-of-configuration (AOC) approximation introduced by Desclaux, Moser & Verhaegen (1917) and Desclaux (1973), expands a non-relativistic *LS* configuration in terms of a number of appropriate *jj* configurations, each taken with a weight that is proportional to the degeneracy of each configuration (Desclaux, 1973; Zatsarinny & Froese Fischer, 2016). The resulting configuration state functions (CSFs) and the weights for all ions are listed in Table S2. For example, Mn^{+} ([Ar] $3d^5\ 4s^1$, isolectronic with chromium) has five CSFs:

$$3d_{3/2}^4\ 3d_{5/2}^1\ 4s^1 \qquad w = 0.02381$$

$$3d_{3/2}^3\ 3d_{5/2}^2\ 4s^1 \qquad w = 0.23810$$

$$3d_{3/2}^2\ 3d_{5/2}^3\ 4s^1 \qquad w = 0.47619$$

$$3d_{3/2}^1\ 3d_{5/2}^4\ 4s^1 \qquad w = 0.23810$$

$$3d_{5/2}^5\ 4s^1 \qquad w = 0.02381$$

where the quantum number j ($j = l \pm s$) is given in the subscript of the orbital symbol and w is the statistical weight of CSF. The Mn^{5+} cation with the ground state *LS*-configuration [Ar] $3d^2$ has three CSFs:

$3d^2_{3/2}$ $w = 0.13333$

$3d^1_{3/2}\ 3d^1_{5/2}$ $w = 0.53333$

$3d^2_{5/2}$ $w = 0.33333$

, while Mn^{7+}, which is isolelectronic with argon ([Ne] $3s^2\ 3p^6$), has only one CSF,

$3s^2\ 3p^2_{1/2}\ 3p^4_{3/2}$ $w = 1.00000$

, because all subshells are filled.

3.2.2 X-ray scattering factors calculations and interpolations

The X-ray scattering factors, and interpolating functions calculations were performed in the Fortran program SF, described in detail in the previous publication (Paper I), which was slightly modified to work with ions. The integration of the scattering factors was performed using the adaptive integrator QAG (Piessens & de Doncker, 1980) from the SLATEC package (Vandevender & Haskell, 1982; Fong et al., 1993) which, when combined with a fine B-spline grid in DSBR_HF integrates the scattering factors with a precision of at least eight decimal digits.

As in Paper I, we employed the $\sin\theta/\lambda$ grid proposed by Wang et al. (1996) which is also used for neutral atoms as given in Table 6.1.1.1 of volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). As discussed above, such a grid is somewhat different from the $\sin\theta/\lambda$ grid used for ions in Table 6.1.1.3 of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006). The relativistic Dirac-Slater X-ray scattering factors of Cromer & Waber (1968b) are truncated at 2.00 $Å^{-1}$ while the non-relativistic Hartree-Fock values of Cromer & Mann (1968a) values terminate at 1.50 $Å^{-1}$. Beyond those cut-off $\sin\theta/\lambda$ values, the scattering factors for the neutral atoms are supposed to be used.

Following the same procedure employed in Paper I for neutral atoms, the function (50) with $m = 4$ and $m = 5$ was used for interpolation of the scattering factors in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} interval, while the expansion (51) of Fox, O'Keefe & Tabbernor (1989) with $n = 3$ and $n = 4$ was utilized in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range.

The starting values for the parameters in the functions (50) and (51) were formed by combining the literature values (Maslen, Fox & O'Keefe, 2006) with those obtained via exhaustive scans of the parameter space in our SF program. The latter were performed on the "VOLTRON" cluster in the MTSU Department of Chemistry that utilizes the AMD Opteron 6378 and 6348 computing processors.

3.3 Assessment of the quality of the calculations

3.3.1 Total electronic energies

Since much less data are available in literature for the relativistic *wavefunctions* of ions than for neutral atoms, the assessment of the quality of the calculated wavefunctions included the total electronic energies and the ionization energies only.

Table S3 lists the total electronic energies for all the species calculated in this work along with the values (where available) from Macchi & Coppens (2001), Rodrigues et al. (2004), and Wang et al. (1996), while the summary of the statistics for the 51 ions common to the first three studies are given in Table 1. We have excluded the Wang et al. (1996) values from the calculation of statistics as it shares only two common ions (Na^+ and Mg^{2+}) with all the other studies.

The discrepancy between our energies and those from Rodrigues et al. (2004) increases steadily with the increase in the atomic number with our values always being lower (more negative). For example, the for B^{2+} (Z=5) the energy difference is just 0.011 hartree (0.04%), which then increases to about 7.8 hartrees (0.12%) for Sb^{3+} and Sb^{5+} (Z=51), and reaches a massve 118 hartrees (0.3%) for Hs^{8+} (Z=108). This is likely due to different levels of

approximation used in this work and Rodrigues et al. (2004). For example, the latter includes several additional QED corrections (Uelhing, Wichmann & Kroll and Källén & Sabry contributions) missing in our work. Rodrigues et al. (2004) also employed the uniform charge nuclear model for the ions with $Z \leq 42$, while all the remaining species in their work and all ions in our study were calculated with the Fermi model.

The differences in the total electronic energies between our work and Macchi & Coppens (2001) are in general smaller as compared to the values of Rodrigues et al. (2004). The Wang et al. (1996) values are similar to those from Macchi & Coppens (2001) and show similar yet slightly larger deviations from our data. Table 1 lists the maximum and mean discrepancies for the energies obtained by Macchi & Coppens (2001) and Rodrigues et al. (2004) relative to our values for the 51 cations common to all three studies, which range from Na^+ to Sb^{5+}.

The largest energy difference for these species between our values and those from Macchi & Coppens (2001) is 7.75 hartree (~0.26%) for Rb^+. For comparison, the discrepancy in energy for Rb^+ between this work and Rodrigues et al. (2004) is 2.60 hartree (0.09%).

For Rodrigues et al. (2004), the largest difference of 7.82 hartrees (0.12%) is found for Sb^{3+} and Sb^{5+} (Z=51) which is not unexpected as these are the two heaviest species in the group of 51 common cations. Indeed, for the next heaviest species, the ions of Te (Z=52), the difference increases to about 8.3 hartrees. However, it is impressive that for the same two Sb cations, the agreement between our work and Macchi & Coppens (2001) is ~0.002 hartree.

The mean absolute energy difference calculated for the common group of 51 ions between our work and the previous studies is noticeably smaller for the Macchi & Coppens (2001) values (0.7±1.4 hartree) as compared to those from Rodrigues et al. (2004) data (2.5±2.5 hartree). However, the differences in energies between our work and Macchi & Coppens

(2001) may seem sporadic at first glance. For example, for Ti^{2+} and Ti^{3+}, our values differ from those from Macchi & Coppens (2001) by about 0.2 hartree, while for Ti^{4+} the energies are essentially identical. For Ag^{+}, the difference is about 0.002 hartree, whereas for Ag^{2+} it increases to 3.2 hartree. Both studies use the Breit and vacuum polarization corrections, and the Fermi nuclear model, and such differences can not be explained by (likely) the lack of self-energy correction in the Macchi & Coppens (2001) study. However, the Macchi & Coppens' (2001) calculations were performed at the optimal level (OL) while we employed the extended average level (EAL). Since the OL method optimizes a single energy level, while the EAL technique works with an average energy of a set of CSFs, the energy (and wavefunction!) differences are expected to be pronounced for states with several CSFs. Indeed, each of the Ti^{2+} and Ti^{3+} cations has several CSFs (Table S2):

Ti^{2+} ([Ar] $3d^2$) CSFs:		Ti^{3+} ([Ar] $3d^1$) CSFs:	
$3d_{3/2}^2$	$w = 0.13333$	$3d_{3/2}^1$	$w = 0.40000$
$3d_{3/2}^1\ 3d_{5/2}^1$	$w = 0.53333$	$3d_{5/2}^1$	$w = 0.60000$
$3d_{5/2}^2$	$w = 0.33333$		

while, Ti^{4+} is isoelectronic with Ar, a closed-shell configuration, and thus has only one CSF. The same explanation applies to the two ions of silver (Table S2):

Ag^{+} ([Kr] $4d^{10}$) CSFs:		Ag^{2+} ([Kr] $4d^9$) CSFs:	
$4d_{3/2}^4\ 4d_{5/2}^6$	$w = 1.00000$	$4d_{3/2}^4\ 4d_{5/2}^5$	$w = 0.60000$
		$4d_{3/2}^3\ 4d_{5/2}^6$	$w = 0.40000$

As such, our energies are expected to show a good agreement with the values of Macchi & Coppens (2001) for the ions with a single CSF, and some discrepancies for the species with multiple CSFs. This explanation applies to essentially all significant energy differences

between our work and Macchi & Coppens (2001) including the valence states of carbon (C_{val}) and silicon (Si_{val}). The discrepancies for the two latter species between our work and Want et al. (1996) are explained analogously. That said, it is unexpected to see a very large difference of 7.75 hartree for Rb^+ which has a closed-shell electronic configuration of krypton. Considering that for Sr^{2+} (which is isoelectronic with Rb^+ and Kr), the agreement between the two calculations is below 0.001 hartree, we suspect either a typo in the reported energy for Rb^+ in Macchi & Coppens (2001) or an error in their calculation (the comparison of the scattering factors for Rb^+ suggests the latter).

3.3.2 Ionization energies

The total electronic energies for a series of cations of an element can be used to estimate the ionization energies. The first atomic ionization energy, I_1, is defined as a difference between the total electronic energy of a neutral species, X, and that of its monovalent cation, X^+:

$$I_1 = E(X^+) - E(X) \tag{52}$$

The second ionization energy is difference in energy between the monovalent and divalent cations:

$$I_2 = E(X^{2+}) - E(X^+) \tag{53}$$

By the same token we define the third, I_3, and higher atomic ionization energies up to I_8

$$I_3 = E(X^{3+}) - E(X^{2+}) \tag{54}$$

$$I_4 = E(X^{4+}) - E(X^{3+})$$

……

$$I_8 = E(X^{8+}) - E(X^{7+})$$

For chemically important ions, the latter is only relevant to ruthenium (Z=45) and osmium (Z=76) that can exist in all oxidation states between +1 and +8 (Greenwood & Earnshaw, 1997). All ionization energies are positive as should be for an endothermic quantity, i.e. the energy of X^{m+} is lower (more negative) than that of X^{n+} for $m < n$.

In Figure S1 and Table S4 we compare the ionization energies ($I_1 \ldots I_8$, in electronvolt, eV) calculated in this work with those from a relatively recent theoretical study by Rodrigues et al. (2004) and *NIST Atomic Spectra Database* (NASD) (Kramida et al., 2021). The latter contains up-to-date values (and their uncertainties) from the most accurate and precise experimental and theoretical studies. That said, for many elements, the higher ionization energies have been estimated from theoretical studies only, including the work by Rodrigues et al. (2004). In Table 2 we summarize the agreement for each ionization energy in terms of the maximum and mean absolute differences relative to the NASD values. For consistency, the NASD entries taken from Rodrigues et al. (2004) have been excluded from the analysis. Figure S1 and Tables 2 and S4 also include several entries calculated from the Macchi & Coppens (2001) data though they are limited to I_3 and I_4. Figure S1a also includes the first ionization energies for Li and Na calculated from the Wang et al. (1996) data. Unfortunately, we could not obtain meaningful I_1 values from the Macchi & Coppens (2001) study by combining their cation data with the neutral atom data of Su & Coppens (1997), and the only two second ionization energies (I_2) calculated from their data for Cu and Ag (34.4 eV and 106.9 eV, respectively) differ significantly from the NASD data (20.3 eV and 21.5 eV, respectively).

Overall, the ionization energies obtained in this work show very similar deviations from the benchmark NASD data as those of Rodrigues et al. (2004), though for higher ionization energies their values are slightly more accurate which is likely a consequence of including additional QED corrections that are missing in our work (see above). Both theoretical studies

pretty much always underestimate the magnitude of the ionization energy, often significantly. Surprisingly, both works show very large discrepancies of above 10 eV for I_6 and I_7 of Pu (Z=94) and I_7 of Np (Z=93) relative to the NASD values taken from an old theoretical study (Carlson et al., 1970), though the estimated uncertainties for the ionization energies of these species listed in NASD are also very large (4-5 eV). For clarity, these entries have been excluded from the calculation of statistics presented in Table 2, though they are listed in Table S4 and Figure S1. For the remaining entries, the largest deviations relative to the NASD data is -7.l eV for this work and -5.7 eV for Rodrigues et al. (2004), both for I_4 of U (Z=92, $I_{4,\mathrm{NASD}}$ = 36.7 eV). For the Macchi & Coppens (2001) study, excluding unreasonably large differences for I_1 and I_2, the largest discrepancy relative to the NASD data is -7.8 eV for I_4 of Ti (Z=22, $I_{4,\mathrm{NASD}}$ = 43.3 eV). For comparison the I_4(Ti) values from this work and Rodrigues et al. (2004) show much smaller discrepancies with the NASD data of 1.1 and 1.3 eV, respectively.

In general, the ionization energies from this work and Rodrigues et al. (2004) agree with each slightly better in comparison to the agreement of each study with the NASD data, which is not surprising as both these studies are theoretical though Rodrigues et al. (2004) included more extensive QED corrections. The Macchi & Coppens (2001) work provides a very limited sample of the ionization energies and shows several unreasonably large deviations relative to the NASD benchmark data. It is interesting to note that the first ionization energy of Li is predicted at 5.34 eV by this work and the Macchi & Coppens (2001) and Wang et al. (1996) calculations which is very close to the literature value of 5.39 eV. For the first ionization energy of Na, this work and Wang et al. (1996) predict essentially identical value of 4.957 eV which is close to 5 eV from Rodrigues et al. (2004) and the experimental value of 5.14 eV. Overall, the analysis suggests that the quality of our relativistic calculations is on approximately the same level as those of Rodrigues et al. (2004) which, when combined with

the previous results obtained for the wavefunctions of neutral atoms (Olukayode, Froese Fischer & Volkov, 2022), instils confidence in the calculated ionic wavefunctions.

3.3.3 Local and integrated wavefunction properties, and orbital energies

For the sake of completeness and as a reference for future studies, we include in Supporting Information a number of tables with

1) the orbital (spinor) energies (core - Table S5, valence - Table S6),
2) the orbital (spinor) mean radii (core – Tables S7, valence – Table S8),
3) the orbital (spinor) charge density maxima (core – Tables S9, valence – Table S10), and
4) the atomic mean radii and mean spherical radii (Table S11).

All these quantities were defined and discussed for the neutral atoms in Paper I.

3.4 X-ray scattering factors and interpolations

As in the previous study, we present separately the discussions of

a) the X-ray scattering factors (section 4.1),
b) the analytical interpolations in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range (section 4.2), and
c) the analytical interpolations in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range (section 4.3).

3.4.1 X-ray scattering factors

The X-ray scattering factors for all 318 ions calculated in this work are given in Table S12 using the $\sin\theta/\lambda$ grid proposed by Wang et al. (1996) and used for neutral atoms in Table 6.1.1.1 of volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006).

A comparison of the X-ray scattering factors for 113 ions common to this study and the previous studies (Table 4) is shown in Figure S2. These include 112 ions (excluding H^-) listed in Table 6.1.1.3 of volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) as well as Cr^{4+} calculated by Rez, Rez & Grant (1994). Note

that because a different set of ions was processed in each of the previous studies, many graphs include data from one or two studies only. Also, while the Doyle & Turner (1968), Thakkar & Smith (1992), Rez, Rez & Grant (1994), Wang at el. (1996) and Macchi & Coppens (2001) studies cover the entire $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range, the Cromer & Waber (1968) and Cromer & Mann (1968) data extend to 2 Å^{-1} only. For the ions calculated in those two studies, the X-ray scattering factor data for $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} were obtained from either Doyle & Turner (1968) or Fox, O'Keefe & Tabbernor (1989). We note that unlike the Doyle & Turner (1968) scattering factor data that correspond to an ion or a neutral atom, the Fox, O'Keefe & Tabbernor (1989) values were obtained via extrapolation of the Cromer & Waber (1968) which, unfortunately, introduces another layer of complexity to the discussion.

We also note that because the Wang at el. (1996) X-ray scattering factors are essentially identical to those from the Macchi & Coppens (2001) study, but the latter includes many more species, we will not provide a separate discussion for the Wang at el. (1996) data.

The results presented in Figure S2 are summarized in Figures 1 – 4. The maximum structure factor differences relative to values obtained in this work are shown in Figure 1a. It is pleasing to see a very good agreement between the values obtained in this work, and the Doyle & Turner (1968) and Thakkar & Smith (1992) data. Except for anions and several heavy ions such as Tb^{3+} and Lu^{3+}, our scattering factors also agree well with the Rez, Rez & Grant (1994) values. The discrepancies for anions in the Rez, Rez & Grant (1994) study are very likely due to the Watson sphere approximation (Watson, 1958) which was also pointed out by Macchi & Coppens (2001). For many species, the agreement is also very good with the Macchi & Coppens (2001) results. In fact, a careful inspection of plots in Figure S2 shows that for all atoms the two studies provide essentially the same scattering factors values for $\sin\theta/\lambda \geq 0.8$ Å^{-1}. The noticeable discrepancies for a number of species (for example, C_{val}, Si_{val}, Cr^{2+}, Mn^{2+}, Fe^{3+} etc.) are observed below 0.8 Å^{-1} (Figure S2) which indicates that it is a valence

electron density effect. These discrepancies are evidently due to the differences between the EAL method used in our study and the OL method employed in Macchi & Coppens (2001). As discussed in section 3.1, the EAL and OL methods are expected to give slightly different results for the configurations with several CSFs. For example, both the C_{val} and Si_{val} configurations include three CSFs while Mn^{2+} ([Ar] $3d^5$) includes five CSFs (Table S2). In comparison, the number of CSFs in Mn^{4+} ([Ar] $3d^3$) is reduced to four, which in turn improves the agreement between the two studies (Figure S2-26). As mentioned in Paper I, from our point of view, the EAL technique provides a more realistic description of the electron distribution for the ground state of an atom/ion in a crystal because it is very unlikely that even an unbonded atom/ion in a crystal is always going to be confined to the lowest energy level of the ground state configuration which is what is optimized in the OL method. For the remaining majority of the ions, the agreement between our work and Macchi & Coppens (2001) is very good. Unfortunately, this is not the case for the non-relativistic Hartree-Fock (Cromer & Mann, 1968) and the relativistic Dirac-Slater (Cromer & Waber, 1968) scattering factors, which show pronounced discrepancies with our values. The deviations for non-relativistic Hartree-Fock values increase with the increase in the atomic number Z which is clearly seen for the Sc^{3+} (Z=21) – Ge^{4+} (Z=32) sequence in Figures 1a and S2. As such, switching to the relativistic Dirac-Slater approximation at Y^{3+} (Z=39), as was done by Cromer & Waber (1968), is certainly warranted. However, as Figures 1a and S2 show, the discrepancies between the Dirac-Slater (Cromer & Waber, 1968) and Dirac-Hartree-Fock (this work) results are quite noticeable in both the low $\sin\theta/\lambda$ region and in the high $\sin\theta/\lambda$ range when for the latter, the Fox, O'Keefe & Tabbernor (1989) extrapolations are used. Finally, we note that for Si_{val} (Z=14) for which the non-relativistic Hartree-Fock approximation may still be valid, our scattering factors fall in between the non-relativistic scattering factors of Cromer & Mann (1968) and relativistic

OL-based Hartree-Fock values of Macchi & Coppens (2001) and Wang et al. (1996) thus providing a reasonable compromise (Figure S2-9).

The mean $|\Delta f|$ differences averaged all $\sin\theta/\lambda$ grid points (Rez, Rez & Grant, 1994) for each atom are plotted in Figure 1b. While these are several times smaller than the maximum $|\Delta f|$ values, they show essentially the same trends as were observed in Figure 1a: the Doyle & Turner (1968), Thakkar & Smith (1992), and the Rez, Rez & Grant (1994) data for the cations agree with our values very well. The scattering factors for anions from the Rez, Rez & Grant (1994) show noticeable deviations due to the employed Watson sphere approximation. The Macchi and Coppens (2001) values are also in a good agreement with our data for the majority of the species, excluding those with a large number of CSFs. The deviations for the non-relativistic Cromer & Mann (1968) values become pronounced at approximately Z=21 (Sc) and increase with Z. The relativistic Dirac-Slater data of Cromer & Mann (1968) show significant deviations for a number of species, especially when augmented in the $\sin\theta/\lambda > 2$ Å^{-1} range with the extrapolations for the neutral atoms by Fox, O'Keefe & Tabbernor (1989) instead of the neutral atom data of Doyle & Turner (1968).

The largest maximum percentile deviations $\Delta f_{\%}$ (Figure 2a) from this work are observed for the non-relativistic Hartree-Fock data for ions starting with Sc^{3+} (0.7 – 1%) and the Dirac-Slater results for a number of species (up to ~2% in Nb^{3+} and Nb^{5+}). The latter are attributed primarily to the issues with extrapolations of the neutral atom data in Fox, O'Keefe & Tabbernor (1989) that complement the Dirac-Slater data of Cromer & Mann (1968) at $\sin\theta/\lambda > 2$ Å^{-1}. The agreement between this work and a) Doyle & Turner (1968), b) Thakkar & Smith (1992), c) Rez, Rez & Grant (1994) (except for anions), and d) Macchi & Coppens (2001), except for C_{val} and Si_{val}, is very good.

Essentially the same trends are observed for the mean $|\Delta f_{\%}|$ values (Figure 2b). Except for the anions in the Rez, Rez & Grant (1994) study, and C_{val} and Si_{val} in Macchi & Coppens (2001), the three relativistic Dirac-Hartree-Fock studies (Doyle & Turner, 1968; Rez, Rez & Grant, 1994; Macchi & Coppens, 2001) show a noticeably better agreement with our values than the non-relativistic Hartree-Fock (Cromer & Mann, 1968) and relativistic Dirac-Slater (Cromer & Waber, 1968) data. Our values are also in a very good agreement with the non-relativistic data for Li^+ and Be^{2+} from Thakkar & Smith (1992) obtained from correlated wavefunctions. Since these are very light atoms, the relativistic effects are expected to be very small, and because each ion contains only two electrons, the correlation effects should be small as well.

In order to investigate the discrepancies between our work and the previous studies as a function of $\sin\theta/\lambda$, we employed the procedure developed in Paper I. We have identified a set of ions common to our work, the data listed in volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006), and the Rez, Rez & Grant (1994) and Macchi & Coppens (2001) studies, and for each ion we have selected the scattering factors at the $\sin\theta/\lambda$ points common to all four studies which happens to be the grid used by Rez, Rez & Grant (1994). In order to avoid issues due to a limited precision of the reported in the literature scattering factors Li^+ and Be^{2+} those ions were excluded from the analysis which resulted in a group of 27 ions that are listed Table 5. For these ions, using the Rez, Rez & Grant (1994) $\sin\theta/\lambda$ grid we have evaluated both

a) the *maximum* Δf and $\Delta f_{\%}$ deviations relative to our values (Figures 3a and 4a), and

b) the *mean* Δf and $\Delta f_{\%}$ deviations relative to our work averaged over the given group of ions at each $\sin\theta/\lambda$ point (Figures 3b and 4b).

In terms of the maximum Δf values (Figure 3a), the largest discrepancies between this work and the Rez, Rez & Grant (1994) and Macchi & Coppens (2001) studies are limited to below approximately 0.5 $Å^{-1}$. In the Rez, Rez & Grant (1994) data these are observed, not unexpectedly, for the F^- and I^- anions due to the use of the Watson sphere approximation. For the Macchi & Coppens (2001) work, the largest discrepancies are found for Mn^{2+} due to conceptual differences between the EAL and OL methods employed in this work and Macchi & Coppens (2001), respectively. Because the Watson sphere and EAL/OL techniques affect primarily the valence charge density, the impact on the core density is very small which in turn results in essentially the same scattering factors in the high $\sin\theta/\lambda$ region. The Maslen, Fox & O'Keefe (2006) data show more pronounced maximum Δf deviations in a wider $\sin\theta/\lambda$ range (0.1 – 2 $Å^{-1}$) than the other two studies which is attributed to using either the Dirac-Slater (Cd^{2+}) or non-relativistic Hartree-Fock (V^{5+} and Cu^{2+}) methods. Beyond 2 $Å^{-1}$, the agreement for many species improves significantly (see, for example, Figure S2-20 for V^{5+}, S2-34 for Cu^{2+} and S2-56 for Cd^{2+}) because the Maslen, Fox & O'Keefe (2006) data switch to the relativistic Hartree-Fock values for neutral atoms from Doyle & Turner (1968). For the ions, for which the Doyle & Turner (1968) neutral atom data for $\sin\theta/\lambda > 2$ $Å^{-1}$ are not available, the Maslen, Fox & O'Keefe (2006) values are taken from extrapolations by Fox, O'Keefe & Tabbernor (1989) which, as discussed in Paper I, display issues for several atoms, including niobium (Z=41) which explains the discrepancies for Nb^{5+} shown in Figure 3a.

The analysis of the mean $|\Delta f|$ discrepancies averaged over 27 common ions (Table 4) at each $\sin\theta/\lambda$ point (Figure 3b) reinforces the conclusions listed above. There is an excellent agreement between this work and the Macchi & Coppens (2001) study at $\sin\theta/\lambda \geq 0.6$ $Å^{-1}$, while the discrepancies below 0.6 $Å^{-1}$ are due to the differences between the EAL and OL approaches. The Rez, Rez & Grant (1994) data show a good agreement with our study at $\sin\theta/\lambda > 0.3$ $Å^{-1}$, though not as good as Macchi & Coppens (2001). The significant

discrepancies for the Rez, Rez & Grant (1994) values below 0.35 Å^{-1} are primarily due to the four anions (F^-, Cl^-, Br^- and I^-) calculated using the Watson method (Watson, 1958) in their study. The Watson approximation introduces a bias in the valence electron density distribution which in turn affect the structure factors in the low $\sin\theta/\lambda$ region, which was also pointed out by Macchi & Coppens (2001). Noticeable discrepancies in the Maslen, Fox & O'Keefe (2006) structure factors are observed for the most important $\sin\theta/\lambda$ range of 0.1 – 2 Å^{-1} which is a consequence of using the non-relativistic Hartree-Fock (Cromer & Mann, 1968) and relativistic Dirac-Slater (Cromer & Waber, 1968) approaches. The agreement improves at $\sin\theta/\lambda > 2$ Å^{-1} due to the use of the Doyle & Turner (1968) values for neutral atoms which are very close to our data for ions. However, for the ions that were not calculated by Doyle & Turner (1968) the discrepancies are due to the use of the extrapolated values of Fox, O'Keefe & Tabbernor (1989).

When the maximum percentile $\Delta f_{\%}$ discrepancies are considered (Figure 4a), the agreement for the 27 common ions between our scattering factors and those from the previous studies is well within 1.5% except for Nb^{5+} at 4.0 Å^{-1}. The scattering factor for Nb^{5+} from our study is 2.460 while the value from Maslen, Fox & O'Keefe (2006) taken from an extrapolation for a neutral atom by Fox, O'Keefe & Tabbernor (1989) is 2.405, which, as pointed out in Paper I, is likely a typo (perhaps, it should have been 2.455). The agreement between our work and Rez, Rez & Grant (1994) and Macchi & Coppens (2001) at $\sin\theta/\lambda \geq 0.4$ Å^{-1} is excellent except for the scattering factors for F^- from Rez, Rez & Grant (1994) at 5 and 6 Å^{-1} which may be a combination of the Watson sphere approximation issues and the limited precision of the reported values. In the low $\sin\theta/\lambda$ region, the maximum $\Delta f_{\%}$ for the Rez, Rez & Grant (1994) study are larger (but still within 1%) than those for the Macchi & Coppens (2001) work, and are observed exclusively for F^-. It suggests that the difference between the EAL/OL methods is not as significant as the use of the Watson sphere approximation. In the $\sin\theta/\lambda \leq 0.4$ Å^{-1}

region, the Maslen, Fox & O'Keefe (2006) discrepancies are very similar in magnitude to or even small than those found for the Macchi & Coppens (2001) study. Because the Maslen, Fox & O'Keefe (2006) data in this $\sin\theta/\lambda$ range come from the non-relativistic Hartree-Fock and relativistic Dirac-Slater wavefunctions, one may argue that a) ignoring the relativistic effects, and b) replacing the Hartree-Fock potential with the Dirac potential is more than satisfactory. However, it would be misleading because i) the ions common to all the studies used in this analysis (Table 4) are not too heavy (Z = 9 – 53), and ii) Figures 1, 2, 3, and S2 clearly show the differences between the relativistic Dirac-Hartree-Fock calculations, and the non-relativistic Hartree-Fock and relativistic Dirac-Slater results. In the $0.4 < \sin\theta/\lambda \leq 2\ \text{Å}^{-1}$ range, the Maslen, Fox & O'Keefe (2006) values show much more pronounced maximum $\Delta f_{\%}$ deviations from our values than the other two studies which reinforces the importance of using the relativistic Hartree-Fock method. A noticeable improvement in the Maslen, Fox & O'Keefe (2006) data at $\sin\theta/\lambda > 2\ \text{Å}^{-1}$ are due to switching to the relativistic Dirac-Hartree-Fock values but the extrapolations used for Nb^{5+}, Zr^{4+} and Pd^{2+} hinder the agreement.

Finally, the analysis of the mean $|\Delta f_{\%}|$ discrepancies averaged over all common ions (Table 5) at each $\sin\theta/\lambda$ grid point (Figure 4b) shows a good overall agreement for all methods as the mean $|\Delta f_{\%}|$ values never exceed 0.2%. Below 0.2 Å^{-1}, our values agree well with the Maslen, Fox & O'Keefe (2006) and Macchi & Coppens (2001) data. At $0.2 < \sin\theta/\lambda \leq 0.4\ \text{Å}^{-1}$, the best agreement is observed with the Rez, Rez & Grant (1994) values. Above 0.4 Å^{-1} the Macchi & Coppens (2001) study shows the smallest discrepancies never exceeding 0.05%. The deviations for the Rez, Rez & Grant (1994) data increase almost linearly from 0.02% at 0.45 Å^{-1} to approximately 0.07 – 0.06% at 3.5 – 4 Å^{-1}, and then go up to 0.14 and 0.13% at 5 and 6 Å^{-1}, respectively, which is due to essentially all halogens' anions (see Figure 4a for F^-, and Figures S2-11, S2-38, and S2-62 for Cl^-, Br^- and I^-, respectively). The

"hybrid" Maslen, Fox & O'Keefe (2006) data show the most significant mean $|\Delta f_{\%}|$ discrepancies at $\sin\theta/\lambda > 0.4$ Å^{-1} which, as mentioned above, is a combination of multiple factors including using the non-relativistic Hartree-Fock (Cromer & Mann, 1968) and relativistic Dirac-Slater (Cromer & Waber, 1968) scattering factors for up to 2 Å^{-1}, and extrapolations for the neutral atoms between 2 and 6 Å^{-1} (Fox, O'Keefe & Tabbernor, 1989).

In summary, the present work not only significantly extends the range of the ions for which the scattering factors are available, but also avoids the issues associated with

a) employing the Watson sphere approximation for monovalent anions,

b) using non-relativistic Hartree-Fock for some atoms and relativistic Dirac-Slater values for others,

c) combining the scattering factors of ions for $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} with those for neutral atoms at $\sin\theta/\lambda > 2$ Å^{-1} especially when the latter were obtained via extrapolations.

3.4.2 Analytical interpolation in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} range

3.4.2.1 Four-Gaussian expansion

The optimized parameters for the interpolating function (50) with $m = 4$ along with the maximum and mean interpolation errors for all 318 calculated species (Table 1) are listed in Table S13 which is analogous to Table 6.1.1.4 in volume C of the 2006 edition of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006).

The maximum and mean errors of the interpolating function (50) with $m = 4$ for all calculated species are also shown with open circles (○) in Figures 5a and 5b, respectively. Note that for convenience a logarithmic scale of base 2 is used in both Figure 5 graphs. In general, there is a significant spread in the magnitude of the maximum and mean errors among the atoms with the largest values being observed for At^- (0.0764 and 0.0194, respectively) followed by At^+ (0.0476 and 0.0126, respectively). A closer inspection of Table S13 and

Figure 5 shows that the largest errors are found for the monovalent anions and monovalent cations, and for most species, the errors decrease with the reduction of the number of electrons. Perhaps, the most representative example is astatine:

	Maximum error:	Mean error:
At^-	0.0764	0.0194
At^+	0.0476	0.0126
At^{3+}	0.0239	0.0080
At^{5+}	0.0075	0.0033
At^{7+}	0.0007	0.0004

These results suggest that the four-Gaussian expansion, equation (50) with $m = 4$, is not flexible enough to provide an accurate description of a relatively steep dependence of the scattering factor on $\sin\theta\,/\lambda$ for anions and low-charge cations. As the number of electrons in a species decreases (thus increasing the charge of the cation), the $f(\sin\theta\,/\lambda)$ curve becomes smoother, which leads to an improvement of the four-Gaussian fit.

For 112 out of 318 the calculated species, it was possible to compare the maximum and mean errors with the values given in Maslen, Fox & O'Keefe (2006). The differences between our values and those reported in literature are plotted in Figure 6. Out of 224 calculated errors (max and mean errors per each atom), in 148 cases the error is lower in this work. The mean improvement is 0.0024 with the most pronounced decreases of 0.0277 and 0.0072 for the maximum and mean interpolating errors, respectively, found for Sb^{5+}. The mean increase in error (worse fit in this work) is 0.0012 which is smaller than the mean improvement. The largest increase in the maximum error is observed for Si_{val} (0.0081), while that for the mean error is found for Cr^{2+} (0.0027). As mentioned in Paper I, the changes in the calculated X-ray scattering factors relative to the published values may indeed lead to small deterioration of the fitting results despite using more sophisticated optimization techniques. However, Figure 6

shows that there are more species for which the interpolating errors, especially the maximum, are smaller in this work as compared to the literature values (Maslen, Fox & O'Keefe, 2006).

3.4.2.2 Five-Gaussian expansion

Increasing the number of Gaussian functions in expansion (50) to five ($m = 5$) results in a significant improvement in the quality of the interpolations. The resulting parameters of the interpolating functions are listed in Table S15, while in Figure 5 the maximum and mean errors of these functions are shown with solid circles (●) for all calculated species. For example, while the largest maximum and mean errors are still observed for At^-, their magnitudes are reduced from 0.0764 to just 0.0156 and from 0.0194 to only 0.0050, respectively. The next largest maximum error at $m = 5$ is 0.0114 in Th^{2+} versus 0.0476 in At^+ at $m = 4$. The second largest mean error is reduced from 0.0126 in At^+ at $m = 4$ to 0.0037 in Rh^+ at $m = 5$.

There are several species for which the maximum error rises slightly as the number of Gaussians in expansion (50) is increased from four to five, however in each case there is an accompanying reduction of the mean error. For example, in Cl^-, the maximum error is increased from 0.0065 to 0.0082 but at the same time, the mean error decreases from 0.0029 to 0.0019. A similar situation is observed for U^{5+}, Pa^{5+}, Np^{5+}, Nb^{2+}, Dy^{3+}, Zr^{2+}, Ni^+, P^+, At^{7+}, Zn^{2+}, Sg^{6+}, Bh^{7+}, U^{6+}, Cf^{4+}, Hs^{8+}, Ho^{3+} and Bk^{4+}. For another group of atoms (Li^+, Be^{2+}, B^{3+}, C^+, C^{4+}, N^{2+}, N^{4+}, N^{5+}, O^+, O^{2+}, Al^{3+}, Si^{4+}, P^{5+}, S^{6+}, Cl^{5+}, Cl^{7+}, Mn^{7+}, Ga^{3+}, Ge^{4+}, As^{5+}, Se^{6+}, Br^{7+}, Os^{8+}, Pu^{4+}, Pu^{5+}, Am^{4+}, Cm^{4+}, Db^{5+}) the changes in the maximum and mean errors upon extending expansion (50) to five Gaussian are negligible.

Overall, out of a total of 636 errors (maximum and minimum errors per each atom), in 590 cases the error is reduced (mean error reduction of 0.0043), for 12 cases there are no changes in the errors (within the used precision of four decimal digits), and in 34 cases the

error is increased with the mean value of 0.0003 which is significantly smaller than the mean error decrease.

These results confirm the superiority of the five-Gaussian interpolating function (50) relative to the expansion with $m = 4$. Because an addition of a single Gaussian function per atom/ion type does not carry a significant computational overhead while producing more accurate scattering factors, the five-Gaussian expansion (50) is highly recommended over the four-Gaussian expansion which is agreement with one of the conclusions in Paper I.

3.4.3 Analytical interpolation in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range

The volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) does not list separate interpolating functions for ions in the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range. Instead, the recommendation is to use the free-atom scattering factors "*because high-angle scattering is dominated by core electrons and is therefore very little affected by ionicity*" (Maslen, Fox & O'Keefe, 2006). We agree with this statement and it is also supported by our data. For example, the scattering factors for the neutral osmium (Paper I) and cations at $\sin\theta/\lambda \geq 2$ Å^{-1} are shown in Table 6. Even when comparing the neutral osmium with Os^{8+}, the differences start in the second decimal digit. Considering that the maximum error of the *four*-term ($n = 3$) expansion (51) for neutral osmium is around 0.19 atomic units, and the mean error is approximately 0.002 atomic units (Paper I) having separate fits for the cations does not look practical. However, upon extending the interpolation function (51) to *five* terms ($n = 4$) the errors reduce to 0.04 and 0.0003 atomic units (Paper I), respectively, which warrants the existence of separate fitting functions for the cations.

3.4.3.1 Four-term expansion

In Table S16 we report parameters of the four-term ($n = 3$) expansions (51) for all 318 species calculated in this work. Because the table mimics Table 6.1.1.4 in volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) the only statistical parameter reported for each fit is the correlation coefficient, C. The quality of our fits for ions can be compared with those for neutral atoms as given in Fox, O'Keefe & Tabbernor (1989) and Maslen, Fox & O'Keefe (2006). In Figure 7a we plot the difference between the correlation coefficients for ions obtained in this work and the correlation coefficients for neutral atoms reported in the literature (Fox, O'Keefe & Tabbernor, 1989) and Maslen, Fox & O'Keefe (2006). Thus, for a series of ions of a given atom, the "literature" value is constant. A positive difference in the correlation coefficient, ΔC, indicates that our fit for an ion is better than the one for a neutral atom reported in the literature. Figure 7a displays many features that are similar to those in Figure 13a in Paper I that included neutral atoms. That is, improvements and reductions in the quality of the fits are observed essentially for the same groups of neutral atoms and ions. For example, better quality fits are found for Cl / K, and Cl^- / K^+, Cu – Br and Cu^+ – Br^-, Dy – W and Dy^{3+} – W^{6+} etc. By the same token, for Ti / Cr and Ti^{2+} – Cr^{3+}, Rb – Mo and Rb^+ – Mo^{6+}, Nd – Eu and Nd^{3+} – Eu^{3+} etc., we see a reduction of the quality of the fit as compared to the literature data. It is not surprising because in this $\sin\theta/\lambda$ range the scattering factors of ions are very similar to those of neutral atoms, and thus the fits must be of a similar quality as well. Overall, out of 112 calculated ΔC differences, a better-quality fit is observed in 61 cases with a mean improvement ΔC of 0.00025, while the quality of the fit was reduced in 50 cases with a mean reduction of C by 0.00012. In the remaining case, the correlation coefficient did not change. The maximum improvement of the fit was observed for Dy^{3+} ($C = 0.999614$) as compared to neutral Dy ($C = 0.9990$) (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006). The maximum reduction of the correlation coefficient

was found for Cr^{2+} ($C = 0.999689$) vs neutral Cr ($C = 1.0000$) (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006). In general, our four-term fits using expansion (51) are of a similar quality or perhaps overall slightly more accurate as compared to the literature data (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006).

In addition to the correlation coefficient, our fitting code also calculates and prints out the maximum and mean errors of the interpolating function (51), the same way it is done for function (51). These are plotted for all 318 species (Table 1) in Figures 8a and 8b, respectively, using open circles (○). The distributions for the two errors for ions follow essentially the same patterns found for neutral atoms in Paper I. The magnitudes of errors for the ions are also very similar to those of neutral atoms (Paper I) which means that the largest errors are found for the ions of the middle of the 5d block. Regrettably, we are still not sure how to explain the observed trends for the interpolating errors.

3.4.3.2 Five-term expansion

As it was found for the neutral atoms (Paper I), extending expansion (51) by one term ($n = 4$) improves the quality of the fit for the majority of the atoms (for many, quite significantly) while for few species the increase in the number of terms does not lead to an improvement. The parameters for the five-term expansion (51) for all 318 calculated species (Table 1) are listed in Table S17. The changes in the correlation coefficient C, ΔC, upon expansion of the interpolating function (51) from $n = 3$ to $n = 4$ are graphed for all calculated species in Figure 7b. Finally, in Figures 8a and 8b the mean and maximum errors of the interpolating function with $n = 4$ (solid circles, ●) are plotted alongside those from the $n = 3$ fits (open circles, ○).

As expected, the curves for $n = 4$ shown in Figures 7b, 8a and 8b display essentially the same features as those for the neutral atoms discussed in Paper I. The most significant

improvements in the correlation coefficient C are observed for the ions of atoms around Sc, Ge/As, Te/I, Ta/W and Bk – Es (Figure 7b). The largest overall positive change for C of 0.000967 is found for Te^{2+} (from 0.998980 at $n = 3$ to 0.999947 at $n = 4$) which is accompanied by a decrease in the maximum/mean errors from 0.17941/0.00149 to 0.03692/0.00035, respectively. Improvements of a similar magnitude are also found for the other cations of Te (Te^{4+}, Te^{5+} and Te^{6+}) as well as ions of iodine (I^{-}, I^{+}, I^{3+}, I^{5+}, and I^{7+}), antimony (Sb^{3+}, Sb^{5+}), scandium (Sc^{+}, Sc^{2+}, Sc^{3+}) and xenon (Xe^{2+}, Xe^{4+}, Xe^{6+} and Xe^{8+}) etc. Essentially no improvement is observed for Be^{2+}, Li^{+}, and the ions of boron, carbon, nitrogen, oxygen, aluminium, silicon, niobium, molybdenum, samarium, astatine, and some others. As discussed in Paper I, the lack of improvement for the lighter elements is due to the fact that even the $n = 3$ expansion interpolates the scattering factors of those species with a high precision. For the heavier elements, such as Nb, Mo, Sm, At, and their ions, we do not currently have a reasonable explanation.

A considerable advantage of the extended expansion (51) is clearly demonstrated in Figures 8a and 8b that show a significant reduction for the majority of the calculated species of both the maximum (Figure 8a) and mean (Figure 8b) errors for the five-term function as compared to the four-term expansion. That said, for a number of species the decrease in the maximum error may lead to slight increase in the mean error (for example, Sm^{2+}, Sm^{3+}, Eu^{2+}, Eu^{3+}, and essentially all ions of Co, Mo and At) and vice versa (cations of Fe, Nb, Nd, Pm, Bi, Po, and Cn). We also note that while the largest mean error of about 0.0014 (Cn^{2+}) is very satisfactory, the largest maximum interpolation errors of 0.10 – 0.19 electrons for many heavy ions despite extending the expansion (51) to five terms may be considered unreasonable. We will rely on feedback from the X-ray crystallography community whether there is a need to improve the accuracy of the interpolating functions before initiating a new study.

3.5 Summary and Concluding remarks

The B-spline Dirac-Hartree-Fock method (Zatsarinny & Froese Fischer, 2016) combined with the extended average level (EAL) approach (Grant et al., 1980; Dyall et al., 1989) was used to re-evaluate fully-relativistic wavefunctions and the corresponding X-ray scattering factors in the $0 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range for all chemically relevant cations and monovalent anions (Greenwood & Earnshaw, 1997) as well as the excied (valence) ns^1np^3 states of the carbon and silicon atoms. From our point of view, this work offers a number of advantages over all the previous studies (Doyle & Turner, 1968; Cromer & Waber, 1968; Cromer & Mann, 1968; Thakkar & Smith, 1992; Rez, Rez & Grant, 1994; Macchi & Coppens, 2001):

1) In comparison to all the earlier studies including those by Rez, Rez & Grant (1994) and Macchi & Coppens (2001), the number of the calculated species has been significantly extended by including all chemically relevant cations and monovalent anions (Greenwood & Earnshaw, 1997).

2) Even though we were unable to compare (due to lack of the available data in literature) the generated wavefunctions for ions the same way it was done for neutral atoms (Olukayode, Froese Fischer & Volkov, 2022), the total electronic energies were subject to a thorough comparison with the previous theoretical studies (Rodrigues et al., 2004 and to a lesser extent, Macchi & Coppens, 2001) in terms of the absolute values and the atomic ionization energies. While our total energies are much closer to the Macchi & Coppens (2001) values, the determined ionization energies match well those of Rodrigues et al. (2004) and show very similar deviations from the experimentally determined quantities (Kramida et al., 2021). A number of local and integrated properties of the calculated wavefunctions that include integrated ionic mean and mean

spherical radii, and locations of the charge density maxima and mean radii of both the core and valence orbitals (spinors) along with their energies, have been deposited as Supplementary Information in order to provide a reference point for future studies.

3) Unlike the X-ray scattering factors presently recommended by IUCr (Maslen, Fox & O'Keefe, 2006) that originate from different sources (Doyle & Turner, 1968; Cromer & Waber, 1968; Cromer & Mann, 1968; Thakkar & Smith, 1992) and were determined at different levels of theory including non-relativistic correlated (Thakkar & Smith, 1992) and Hartree-Fock methods (Cromer & Mann, 1968), and relativistic Dirac-Hartree-Fock (Doyle & Turner, 1968) and Dirac-Slater (Cromer & Waber, 1968) approaches, our values come from a single relativistic Dirac-Hartree-Fock study that includes both the Breit interaction correction and the Fermi nuclear charge model.

4) Just like in our study of the neutral atoms (Olukayode, Froese Fischer & Volkov, 2022) the B-spline representation of the radial functions combined with a dense radial grid allowed for a precise determination of the X-ray scattering factors values thus avoiding ambiguity associated with the numerical integration procedures used in essentially all the previous investigations.

5) In contrast to the recommended approach (Maslen, Fox & O'Keefe, 2006) in which the X-ray scattering factors for ions beyond 2 $Å^{-1}$ (in several instances, above 1.50 $Å^{-1}$) are approximated by the values derived for neutral atoms that for many species were obtained via the extrapolation procedure (Fox, O'Keefe & Tabbernor, 1989), our X-ray scattering factors , just like the Rez, Rez & Grant (1994) and Macchi & Coppens (2001) values, have been obtained directly from the ionic wavefunctions for the entire 0 – 6 $Å^{-1}$ range. The issues with the extrapolation approach were discussed in detail in our previous study (Olukayode, Froese Fischer & Volkov, 2022).

6) While Rez, Rez & Grant (1994) had to resolve to the Watson sphere approximation when calculating anions, the present study did not require that and included a uniform treatment of all cations and monovalent anions.

7) The X-ray scattering factors in both the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} ranges have been interpolated using the recommended functions (50) (Vand, Eiland & Pepinsky, 1957; Doyle & Turner, 1968; Cromer & Waber, 1968; Maslen, Fox & O'Keefe, 2006) and (51) (Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006) at the conventional and extended levels. In comparison, the interpolating functions were not included in the Rez, Rez & Grant (1994) study, while Macchi & Coppens (2001) used the unconventional six-Gaussian expansions in the 0 – 2, 2 – 4, and 4 – 6 Å^{-1} intervals.

8) The presented extended expansions (50) and (51) with $m = 5$ and $n = 4$, respectively, offer a significant improvement in the accuracy of the interpolated values for the X-ray scattering factors while preserving the general mathematical form of the established equations.

9) The generated relativistic Dirac-Hartree-Fock wavefunctions for neutral atoms (Olukayode, Froese Fischer & Volkov, 2022) and ions (this work) stored in the B-spline representations can be easily used to create custom fits of desired accuracy. For example, six-Gaussian fits used in the Su & Coppens (1997, 1998a) and Macchi & Coppens (2001) studies can be readily obtained.

10) A thorough comparison of the X-ray scattering factors of ions obtained in this work with those from all the previous studies allowed for a better understanding of the effects of the average / extended average level (AL/EAL) and optimal level (OL)

approximations used in the relativistic Dirac-Hartree-Fock calculations. As a reference for future studies, we include in the Supplementary Information a complete list of the employed ground state electronic configurations of the ions with the associated configuration state functions (CSFs) and their weights.

The future plans include a) calculation of the relativistic Dirac-Hartree-Fock scattering factors for all chemically relevant *multivalent* anions (Greenwood & Earnshaw, 1997), and b) determination of the analytical representations of the relativistic Dirac-Hartree-Fock wavefunctions for all neutral atoms and ions using a linear combination of Slater-type functions as was done in the studies by Su & Coppens (1998b) and Macchi & Coppens (2001).

Acknowledgements A number of symbolical and numerical computations for debugging purposes were performed in the *Mathematica* system (Wolfram Research, 2022). DBSR_HF and SF calculations for ions were performed on the MTSU VOLTRON cluster supported by the MTSU Chemistry Department, Information Technology Division, and the Computational Science Ph.D. Program. Their continuous support is gratefully acknowledged.

Table 5 The list and numbering of the species (cations, monovalent anions, and the excited (valence) states of carbon and silicon) included in this work.

Numb	Specie	Numb	Specie	Numb	Specie	Numb	Specie	Numb	Specie
1	Li^{+}	41	Cl^{2+}	81	Co^{2+}	121	Zr^{4+}	161	In^{3+}
2	Be^{2+}	42	Cl^{3+}	82	Co^{3+}	122	Nb^{2+}	162	Sn^{2+}
3	B^{+}	43	Cl^{4+}	83	Co^{4+}	123	Nb^{3+}	163	Sn^{4+}
4	B^{2+}	44	Cl^{5+}	84	Co^{5+}	124	Nb^{4+}	164	Sb^{3+}
5	B^{3+}	45	Cl^{6+}	85	Ni^{+}	125	Nb^{5+}	165	Sb^{5+}
6	C^{+}	46	Cl^{7+}	86	Ni^{2+}	126	Mo^{+}	166	Te^{2+}
7	C^{2+}	47	Cl^{-}	87	Ni^{3+}	127	Mo^{2+}	167	Te^{4+}
8	C^{3+}	48	K^{+}	88	Ni^{4+}	128	Mo^{3+}	168	Te^{5+}
9	C^{4+}	49	Ca^{2+}	89	Cu^{+}	129	Mo^{4+}	169	Te^{6+}
10	C_{val}	50	Sc^{+}	90	Cu^{2+}	130	Mo^{5+}	170	I^{+}
11	N^{+}	51	Sc^{2+}	91	Cu^{3+}	131	Mo^{6+}	171	I^{3+}
12	N^{2+}	52	Sc^{3+}	92	Cu^{4+}	132	Tc^{+}	172	I^{5+}
13	N^{3+}	53	Ti^{2+}	93	Zn^{2+}	133	Tc^{2+}	173	I^{7+}
14	N^{4+}	54	Ti^{3+}	94	Ga^{+}	134	Tc^{3+}	174	I^{-}
15	N^{5+}	55	Ti^{4+}	95	Ga^{2+}	135	Tc^{4+}	175	Xe^{2+}
16	O^{+}	56	V^{+}	96	Ga^{3+}	136	Tc^{5+}	176	Xe^{4+}
17	O^{2+}	57	V^{2+}	97	Ge^{+}	137	Tc^{6+}	177	Xe^{6+}
18	O^{-}	58	V^{3+}	98	Ge^{2+}	138	Tc^{7+}	178	Xe^{8+}
19	F^{-}	59	V^{4+}	99	Ge^{3+}	139	Ru^{+}	179	Cs^{+}
20	Na^{+}	60	V^{5+}	100	Ge^{4+}	140	Ru^{2+}	180	Ba^{2+}
21	Mg^{2+}	61	Cr^{+}	101	As^{2+}	141	Ru^{3+}	181	La^{2+}
22	Al^{+}	62	Cr^{2+}	102	As^{3+}	142	Ru^{4+}	182	La^{3+}
23	Al^{3+}	63	Cr^{3+}	103	As^{5+}	143	Ru^{5+}	183	Ce^{2+}
24	Si^{+}	64	Cr^{4+}	104	Se^{2+}	144	Ru^{6+}	184	Ce^{3+}
25	Si^{2+}	65	Cr^{5+}	105	Se^{4+}	145	Ru^{7+}	185	Ce^{4+}
26	Si^{3+}	66	Cr^{6+}	106	Se^{6+}	146	Ru^{8+}	186	Pr^{2+}
27	Si^{4+}	67	Mn^{+}	107	Br^{+}	147	Rh^{+}	187	Pr^{3+}
28	Si_{val}	68	Mn^{2+}	108	Br^{3+}	148	Rh^{2+}	188	Pr^{4+}
29	P^{+}	69	Mn^{3+}	109	Br^{4+}	149	Rh^{3+}	189	Nd^{2+}
30	P^{2+}	70	Mn^{4+}	110	Br^{5+}	150	Rh^{4+}	190	Nd^{3+}
31	P^{3+}	71	Mn^{5+}	111	Br^{7+}	151	Rh^{5+}	191	Pm^{3+}
32	P^{4+}	72	Mn^{6+}	112	Br^{-}	152	Rh^{6+}	192	Sm^{2+}
33	P^{5+}	73	Mn^{7+}	113	Kr^{2+}	153	Pd^{2+}	193	Sm^{3+}
34	S^{+}	74	Fe^{+}	114	Rb^{+}	154	Pd^{4+}	194	Eu^{2+}
35	S^{2+}	75	Fe^{2+}	115	Sr^{2+}	155	Ag^{+}	195	Eu^{3+}
36	S^{3+}	76	Fe^{3+}	116	Y^{2+}	156	Ag^{2+}	196	Gd^{+}
37	S^{4+}	77	Fe^{4+}	117	Y^{3+}	157	Ag^{3+}	197	Gd^{2+}
38	S^{5+}	78	Fe^{5+}	118	Zr^{+}	158	Cd^{2+}	198	Gd^{3+}
39	S^{6+}	79	Fe^{6+}	119	Zr^{2+}	159	In^{+}	199	Tb^{+}
40	Cl^{+}	80	Co^{+}	120	Zr^{3+}	160	In^{2+}	200	Tb^{3+}

Table 6. Continued

Numb	Specie	Numb	Specie	Numb	Specie	Numb	Specie	Numb	Specie
201	Tb^{4+}	226	Re^{3+}	251	Au^{3+}	276	Pa^{4+}	301	Cf^{2+}
202	Dy^{2+}	227	Re^{4+}	252	Au^{5+}	277	Pa^{5+}	302	Cf^{3+}
203	Dy^{3+}	228	Re^{5+}	253	Hg^{+}	278	U^{3+}	303	Cf^{4+}
204	Ho^{3+}	229	Re^{6+}	254	Hg^{2+}	279	U^{4+}	304	Es^{2+}
205	Er^{3+}	230	Re^{7+}	255	Tl^{+}	280	U^{5+}	305	Es^{3+}
206	Tm^{2+}	231	Os^{+}	256	Tl^{3+}	281	U^{6+}	306	Fm^{2+}
207	Tm^{3+}	232	Os^{2+}	257	Pb^{2+}	282	Np^{3+}	307	Fm^{3+}
208	Yb^{2+}	233	Os^{3+}	258	Pb^{4+}	283	Np^{4+}	308	Md^{2+}
209	Yb^{3+}	234	Os^{4+}	259	Bi^{3+}	284	Np^{5+}	309	Md^{3+}
210	Lu^{3+}	235	Os^{5+}	260	Bi^{5+}	285	Np^{6+}	310	No^{2+}
211	Hf^{2+}	236	Os^{6+}	261	Po^{2+}	286	Np^{7+}	311	No^{3+}
212	Hf^{3+}	237	Os^{7+}	262	Po^{4+}	287	Pu^{3+}	312	Lr^{3+}
213	Hf^{4+}	238	Os^{8+}	263	Po^{6+}	288	Pu^{4+}	313	Rf^{4+}
214	Ta^{2+}	239	Ir^{+}	264	At^{+}	289	Pu^{5+}	314	Db^{5+}
215	Ta^{3+}	240	Ir^{2+}	265	At^{3+}	290	Pu^{6+}	315	Sg^{6+}
216	Ta^{4+}	241	Ir^{3+}	266	At^{5+}	291	Pu^{7+}	316	Bh^{7+}
217	Ta^{5+}	242	Ir^{4+}	267	At^{7+}	292	Am^{2+}	317	Hs^{8+}
218	W^{+}	243	Ir^{5+}	268	At^{-}	293	Am^{3+}	318	Cn^{2+}
219	W^{2+}	244	Ir^{6+}	269	Fr^{+}	294	Am^{4+}		
220	W^{3+}	245	Pt^{2+}	270	Ra^{2+}	295	Am^{5+}		
221	W^{4+}	246	Pt^{4+}	271	Ac^{3+}	296	Am^{6+}		
222	W^{5+}	247	Pt^{5+}	272	Th^{2+}	297	Cm^{3+}		
223	W^{6+}	248	Pt^{6+}	273	Th^{3+}	298	Cm^{4+}		
224	Re^{+}	249	Au^{+}	274	Th^{4+}	299	Bk^{3+}		
225	Re^{2+}	250	Au^{2+}	275	Pa^{3+}	300	Bk^{4+}		

Table 6 The mean and maximum differences in the total electronic energies E (atomic units) for the 51 common ions calculated in this work relative to those in Macchi & Coppens (2001) and Rodrigues et al. (2004). A negative maximum ΔE means that the energy in this work is lower. For each maximum value, the ion for which that value is observed is listed in parentheses.

	Rodrigues et al. (2004)	**Macchi & Coppens (2001)**
Mean $\lvert\Delta E\rvert$	2.55	0.70
Standard deviation of the mean $\lvert\Delta E\rvert$	2.48	1.38
Maximum ΔE (ion)	-7.82 (Sb^{3+} / Sb^{5+})	-7.75 (Rb^{+})

Table 7 The mean and maximum differences for the selected ionization energies (eV) calculated in this work, Macchi & Coppens (2001) and Rodrigues et al. (2004) relative to the *NIST Atomic Spectra Database* (NASD) data (Kramida et al., 2021).

	This work	Rodrigues et al. (2004)	Macchi & Coppens (2001)
First ionization energy, I_1			
Number of entries	42	42	—
Mean $\lvert\Delta I_1\rvert$	0.9	1.0	—
Maximum ΔI_1 (atom / $I_{1,\mathrm{NASD}}$)	-1.9 (Hg / 10.4)	-2.4 (Os / 8.4)	—
Second ionization energy, I_2			
Number of entries	32	32	—[1]
Mean $\lvert\Delta I_2\rvert$	1.5	1.3	—[1]
Maximum ΔI_2 (atom / $I_{2,\mathrm{NASD}}$)	-3.0 (Cr / 16.5)	-3.3 (Cu / 20.3)	—[1]
Third ionization energy, I_3			
Number of entries	51	51	7[2]
Mean $\lvert\Delta I_3\rvert$	1.7	1.5	1.7[2]
Maximum ΔI_3 (atom / $I_{3,\mathrm{NASD}}$)	-4.9 (Eu / 24.8)	-5.7 (Tm / 23.7)	-2.6 (Ni/ 35.19)[2]
Fourth ionization energy, I_4			
Number of entries	32[3]	32[3]	4[4]
Mean $\lvert\Delta I_4\rvert$	2.0[3]	1.6[3]	2.5[4]
Maximum ΔI_4 (atom / $I_{4,\mathrm{NASD}}$)	-7.1 (U / 36.7)[3]	-5.7 (U / 36.7)[3]	-7.8 (Ti / 43.3)[4]
Fifth ionization energy, I_5			
Number of entries	15[3]	15[3]	—
Mean $\lvert\Delta I_5\rvert$	1.9[3]	1.1[3]	—
Maximum ΔI_5 (atom / $I_{5,\mathrm{NASD}}$)	-4.1 (Co / 79.5)[3]	-3.3 (Te / 59.3)[3]	—
Sixth ionization energy, I_6			
Number of entries	6[3,5]	6[3,5]	—
Mean $\lvert\Delta I_6\rvert$	1.8[3,5]	1.0[3,5]	—
Maximum ΔI_6 (atom / $I_{6,\mathrm{NASD}}$)	-2.6 (Mn / 95.6)[3,5]	-1.9 (Cl / 96.9)[3,5]	—
Seventh ionization energy, I_7			
Number of entries	3[3,6]	3[3,6]	
Mean $\lvert\Delta I_7\rvert$	1.2[3,6]	1.0[3,6]	—
Maximum ΔI_7 (atom / $I_{7,\mathrm{NASD}}$)	-1.9 (Re / 82.7)[3,6]	-1.2 (Cl / 114.2 and Mn / 119.2)[3,6]	—
Eighth ionization energy, I_8 (Ru and Os only)			
ΔI_8 (Ru)	-0.24	0[7]	
ΔI_8 (Os) ($I_{8,\mathrm{NASD}}$)	-2.0 (102.0)	-1.0 (102.0)	

[1] For the only two species for which I_2 can be calculated using the Macchi & Coppens (2001) data, Cu and Ag, the values (34.4 eV and 106.9 eV, respectively) differ significantly from the NASD data (20.3 eV and 21.5 eV, respectively).

[2] The I_3 values were calculated for only seven species: Ti, V, Cr, Mn, Fe, Co, and Ni.

[3] Because for several species, NASD includes the Rodrigues et al. (2004) data, those values were excluded from the calculation of statistics.

[4] The I_4 values were calculated for only four species: Ti, Mn, Ru, and Rh.

[5] Due to significant discrepancies between the NASD I_6 value for Pu (80 eV) and those from this work (64.7 eV) and Rodrigues et al. (2004) (67 eV) these values have been excluded from the statistics calculation.

[6] Due to significant discrepancies between the NASD I_7 values for Np (92 eV) and Pu (95 eV) and those from this work (80.1 and 82.5 eV, respectively) and Rodrigues et al. (2004) (81 and 84 eV, respectively) these values have been excluded from the statistics calculation.

[7] The NASD value for I_8 of Ru includes the Rodrigues et al. (2004) value.

Table 8 The numbering of 113 species used for the analysis presented in Figures 15, 16, 19 and 22a.

Number	Species	Number	Species	Number	Species	Number	Species
1	Li^{+}	31	Ni^{2+}	61	Sb^{5+}	91	Pt^{4+}
2	Be^{2+}	32	Ni^{3+}	62	I^{-}	92	Au^{+}
3	C_{val}	33	Cu^{+}	63	Cs^{+}	93	Au^{3+}
4	O^{-}	34	Cu^{2+}	64	Ba^{2+}	94	Hg^{+}
5	F^{-}	35	Zn^{2+}	65	La^{3+}	95	Hg^{2+}
6	Na^{+}	36	Ga^{3+}	66	Ce^{3+}	96	Tl^{+}
7	Mg^{2+}	37	Ge^{4+}	67	Ce^{4+}	97	Tl^{3+}
8	Al^{3+}	38	Br^{-}	68	Pr^{3+}	98	Pb^{2+}
9	Si_{val}	39	Rb^{+}	69	Pr^{4+}	99	Pb^{4+}
10	Si^{4+}	40	Sr^{2+}	70	Nd^{3+}	100	Bi^{3+}
11	Cl^{-}	41	Y^{3+}	71	Pm^{3+}	101	Bi^{5+}
12	K^{+}	42	Zr^{4+}	72	Sm^{3+}	102	Ra^{2+}
13	Ca^{2+}	43	Nb^{3+}	73	Eu^{2+}	103	Ac^{3+}
14	Sc^{3+}	44	Nb^{5+}	74	Eu^{3+}	104	Th^{4+}
15	Ti^{2+}	45	Mo^{3+}	75	Gd^{3+}	105	U^{3+}
16	Ti^{3+}	46	Mo^{5+}	76	Tb^{3+}	106	U^{4+}
17	Ti^{4+}	47	Mo^{6+}	77	Dy^{3+}	107	U^{6+}
18	V^{2+}	48	Ru^{3+}	78	Ho^{3+}	108	Np^{3+}
19	V^{3+}	49	Ru^{4+}	79	Er^{3+}	109	Np^{4+}
20	V^{5+}	50	Rh^{3+}	80	Tm^{3+}	110	Np^{6+}
21	Cr^{2+}	51	Rh^{4+}	81	Yb^{2+}	111	Pu^{3+}
22	Cr^{3+}	52	Pd^{2+}	82	Yb^{3+}	112	Pu^{4+}
23	Cr^{4+}	53	Pd^{4+}	83	Lu^{3+}	113	Pu^{6+}
24	Mn^{2+}	54	Ag^{+}	84	Hf^{4+}		
25	Mn^{3+}	55	Ag^{2+}	85	Ta^{5+}		
26	Mn^{4+}	56	Cd^{2+}	86	W^{6+}		
27	Fe^{2+}	57	In^{3+}	87	Os^{4+}		
28	Fe^{3+}	58	Sn^{2+}	88	Ir^{3+}		
29	Co^{2+}	59	Sn^{4+}	89	Ir^{4+}		
30	Co^{3+}	60	Sb^{3+}	90	Pt^{2+}		

Table 9 The list of 27 ions, excluding Li^+ and Be^{2+}, common to this work, and the IUCr (Maslen, Fox & O'Keefe, 2006), Rez, Rez & Grant (1994) and Macchi & Coppens (2001) publications used for the analysis presented in Figures 17 and 18.

Number	Species	Number	Species	Number	Species	Number	Species
1	F^-	8	Sc^{3+}	15	Cu^{2+}	22	Nb^{5+}
2	Na^+	9	Ti^{4+}	16	Zn^{2+}	23	Mo^{6+}
3	Mg^{2+}	10	V^{5+}	17	Br^-	24	Pd^{2+}
4	Al^{3+}	11	Mn^{2+}	18	Rb^+	25	Ag^{2+}
5	Cl^-	12	Fe^{2+}	19	Sr^{2+}	26	Cd^{2+}
6	K^+	13	Co^{2+}	20	Y^{3+}	27	I^-
7	Ca^{2+}	14	Ni^{2+}	21	Zr^{4+}		

Table 10 The X-ray scattering factors of neutral osmium (Olukayode, Froese Fischer & Volkov, 2022) and its cations (this work) for $\sin\theta/\lambda \geq 2$ Å^{-1}. The identical digits (after rounding) are shown in **bold**. The differences are underlined.

$\sin\theta/\lambda$ (Å^{-1})	Os	Os^+	Os^{2+}	Os^{3+}	Os^{4+}	Os^{5+}	Os^{6+}	Os^{7+}	Os^{8+}
2.00	**14.237** 01	**14.237** 30	**14.237** 67	**14.236** 99	**14.236** 10	**14.235** 02	**14.233** 79	**14.232** 43	**14.230** 97
2.50	**11.540** 64	**11.541** 13	**11.541** 76	**11.540** 79	**11.539** 31	**11.537** 29	**11.534** 72	**11.531** 60	**11.527** 92
3.00	**9.1392** 4	**9.1393** 4	**9.1393** 9	**9.1387** 6	**9.1377** 2	**9.1362** 3	**9.1342** 5	**9.1317** 7	**9.1287** 7
3.50	**7.2868** 2	**7.2868** 1	**7.2867** 3	**7.2863** 6	**7.2857** 4	**7.2848** 5	**7.2836** 7	**7.2821** 8	**7.2803** 8
4.00	**6.0566** 6	**6.0566** 8	**6.0566** 9	**6.0565** 0	**6.0562** 0	**6.0557** 8	**6.0552** 5	**6.0545** 8	**6.0537** 9
5.00	**4.7836** 9	**4.7836** 7	**4.7836** 6	**4.7837** 9	**4.7839** 9	**4.7842** 5	**4.7845** 8	**4.7849** 7	**4.7854** 2
6.00	**4.0371** 9	**4.0370** 4	**4.0368** 4	**4.0371** 8	**4.0376** 3	**4.0382** 0	**4.0388** 8	**4.0396** 7	**4.0405** 7

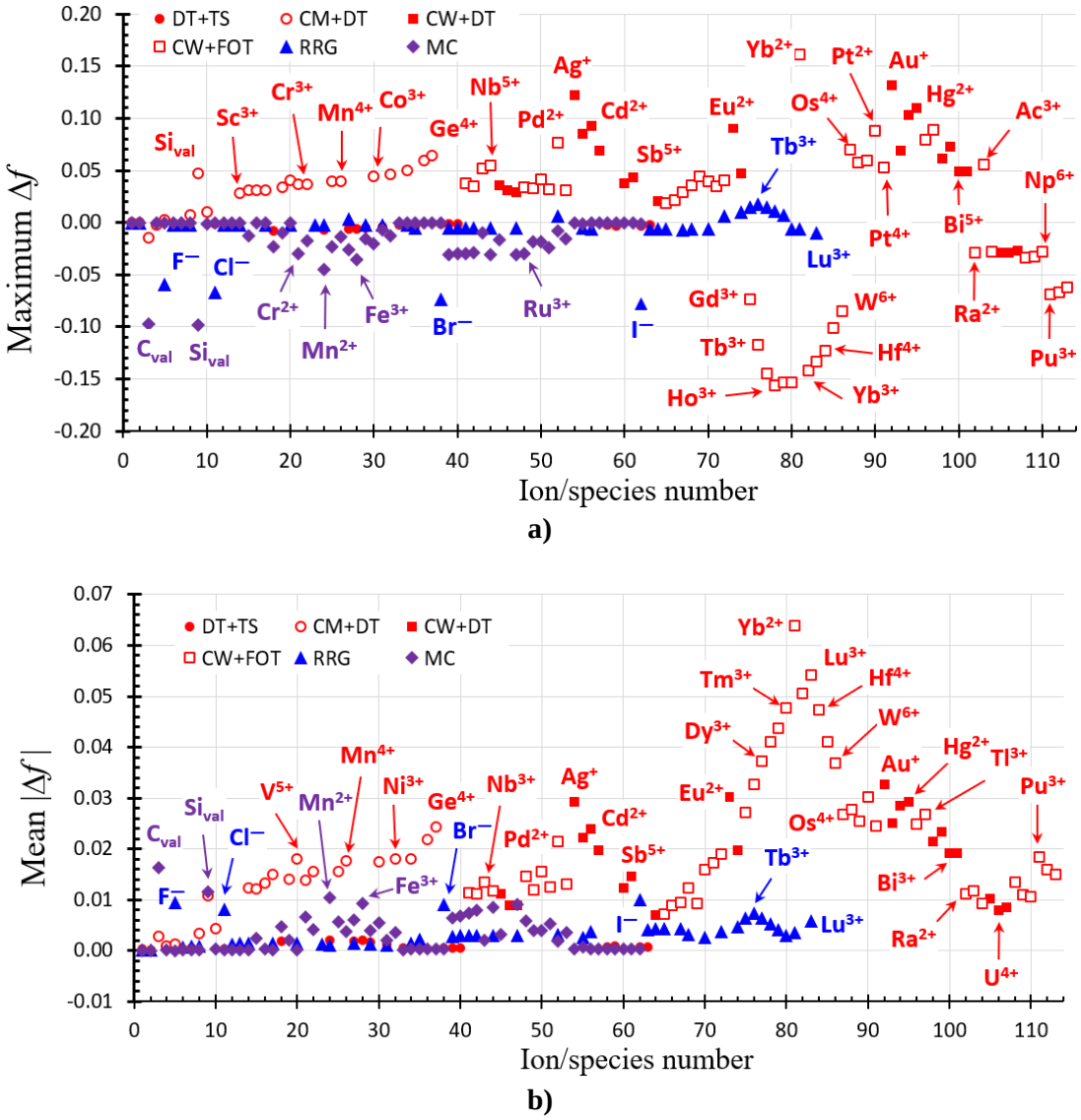


Figure 15 The **(a)** *maximum* and **(b)** *mean differences* in the X-ray scattering factors (Δf) between this work and the previous studies for 113 common species (Table 4) plotted as a function of the atomic number, *Z*. The previous studies are identified as follows: ● DT+TS - Doyle & Turner (1968) and Thakkar & Smith (1992) data, ○ CM+DT - Cromer & Mann (1968) values for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented by the Doyle & Turner (1968) values for $2 \le \sin\theta/\lambda \le 6$ Å^{-1}, ■ CW+DT - Cromer & Waber (1968) data for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented by the Doyle & Turner (1968) values for $2 \le \sin\theta/\lambda \le 6$ Å^{-1}, □ CW+FOT - Cromer & Waber (1968) values for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented for $2 \le \sin\theta/\lambda \le 6$ Å^{-1} with extrapolations by Fox, O'Keefe & Tabbernor (1989); ▲ RRG - Rez, Rez & Grant (1994), and ♦ MC - Macchi & Coppens (2001). The $\sin\theta/\lambda$ grid from the Rez, Rez & Grant (1994) study was used. The averaging in **(b)** was done over all the $\sin\theta/\lambda$ grid points between 0 and 6 Å^{-1}.

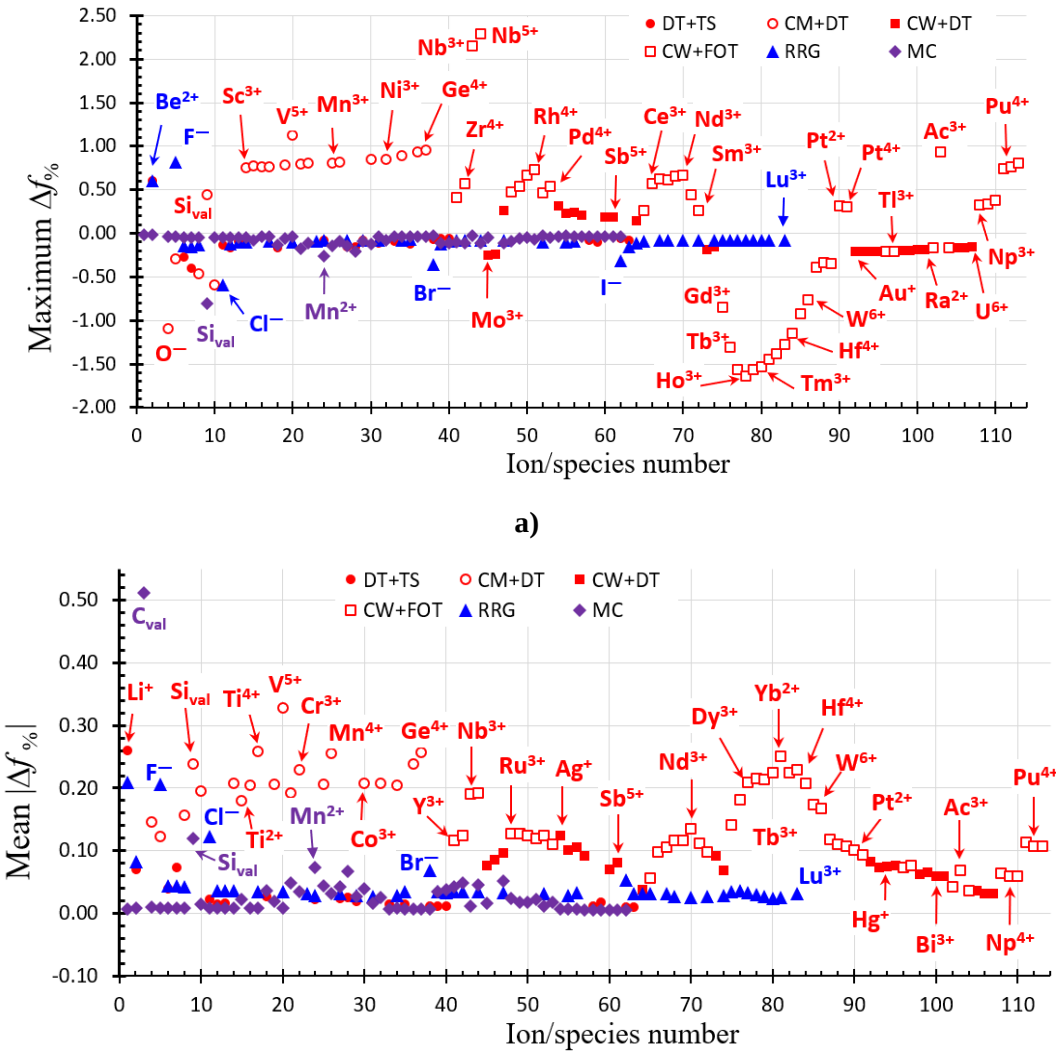


Figure 16 . The **(a)** *maximum* and **(b)** *mean percent differences* in the X-ray scattering factors ($\Delta f_{\%}$) between this work and the previous studies for 113 common ions (Table 4) plotted as a function of the atomic number, *Z*. The previous studies are identified as follows: ● DT+TS - Doyle & Turner (1968) and Thakkar & Smith (1992) data, ○ CM+DT - Cromer & Mann (1968) values for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented by the Doyle & Turner (1968) values for $2 \le \sin\theta/\lambda \le 6$ Å^{-1}, ■ CW+DT - Cromer & Waber (1968) values for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented by the Doyle & Turner (1968) values for $2 \le \sin\theta/\lambda \le 6$ Å^{-1}, □ CW+FOT - Cromer & Waber (1968) values for $0 \le \sin\theta/\lambda \le 2$ Å^{-1} augmented for $2 \le \sin\theta/\lambda \le 6$ Å^{-1} by extrapolations by Fox, O'Keefe & Tabbernor (1989); ▲ RRG - Rez, Rez & Grant (1994), and ♦ MC - Macchi & Coppens (2001). The $\sin\theta/\lambda$ grid from the Rez, Rez & Grant (1994) study was used. The averaging in **(b)** was done over all the $\sin\theta/\lambda$ grid points between 0 and 6 Å^{-1}. Not shown in **(a)** $\Delta f_{\%,\max}$ of -4.2% and -2.5% for C_{val} in the CM+DT and SCM data, respectively, and in **(b)** $|\Delta f_{\%}|_{\text{mean}}$ of 1.0% for C_{val} in the CM+DT data.

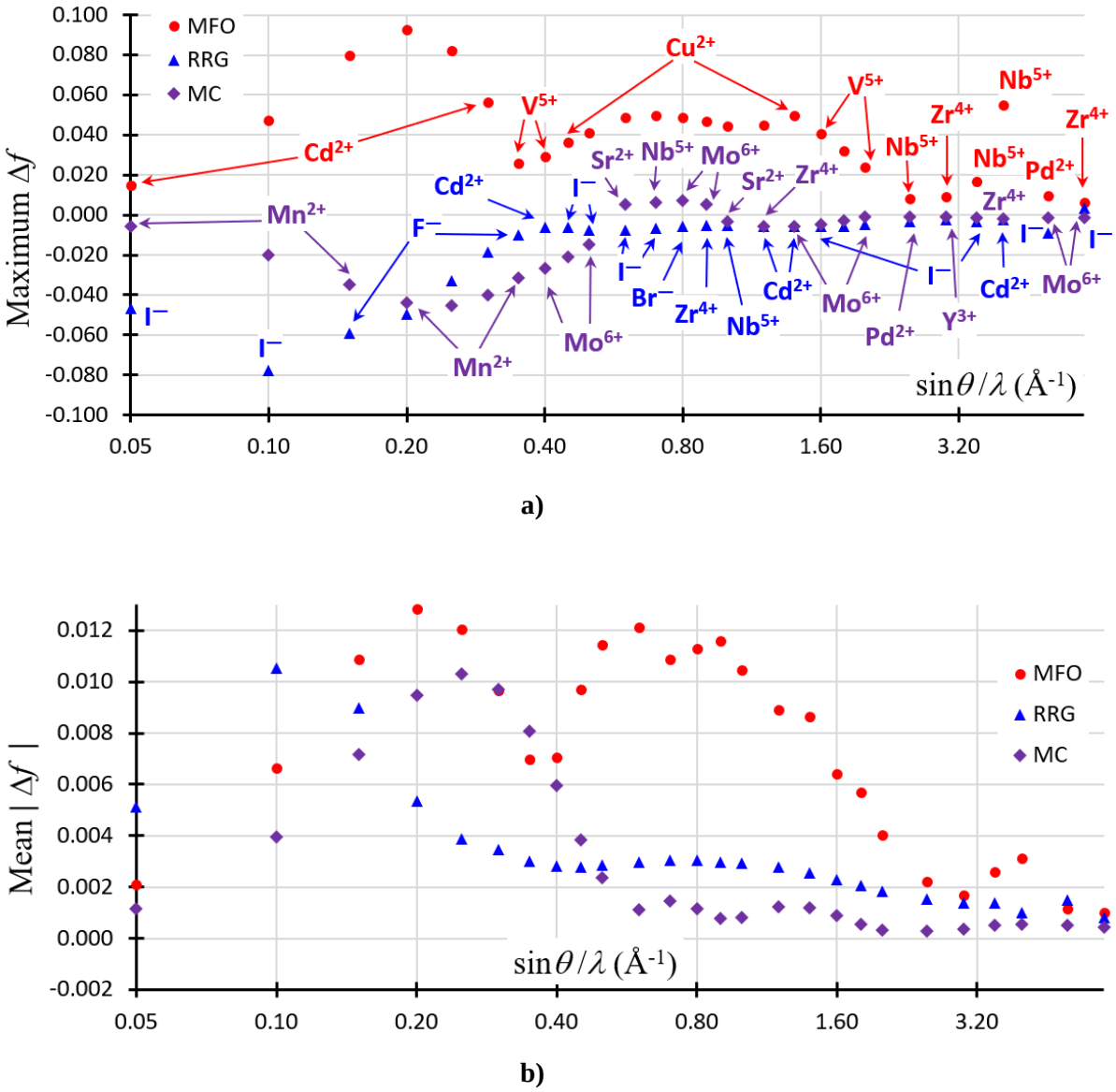
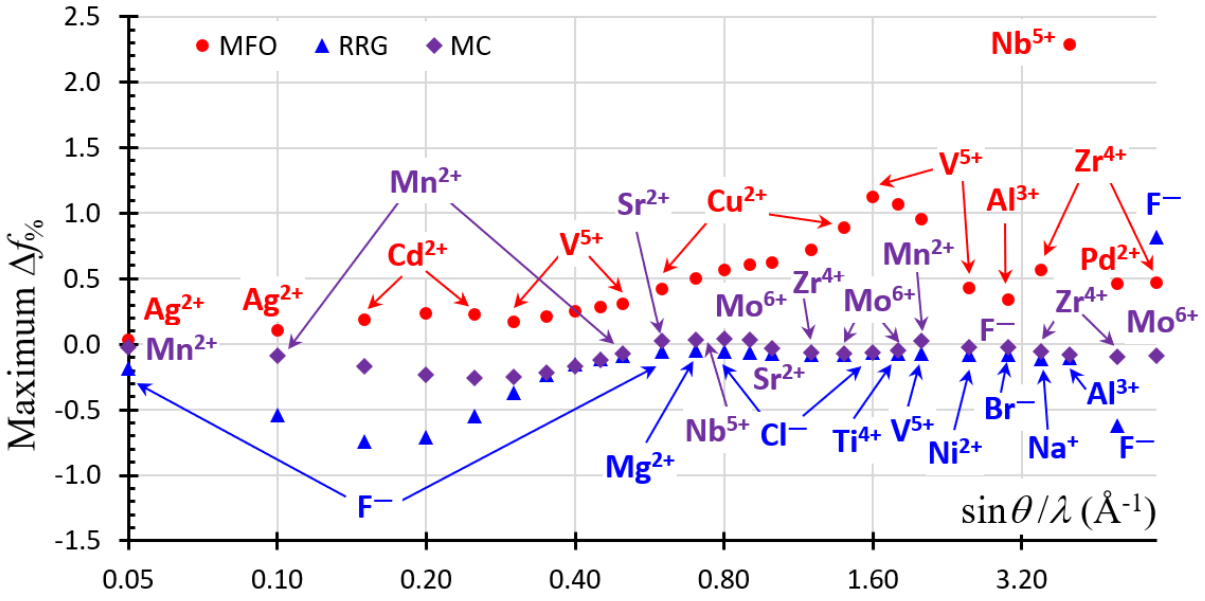
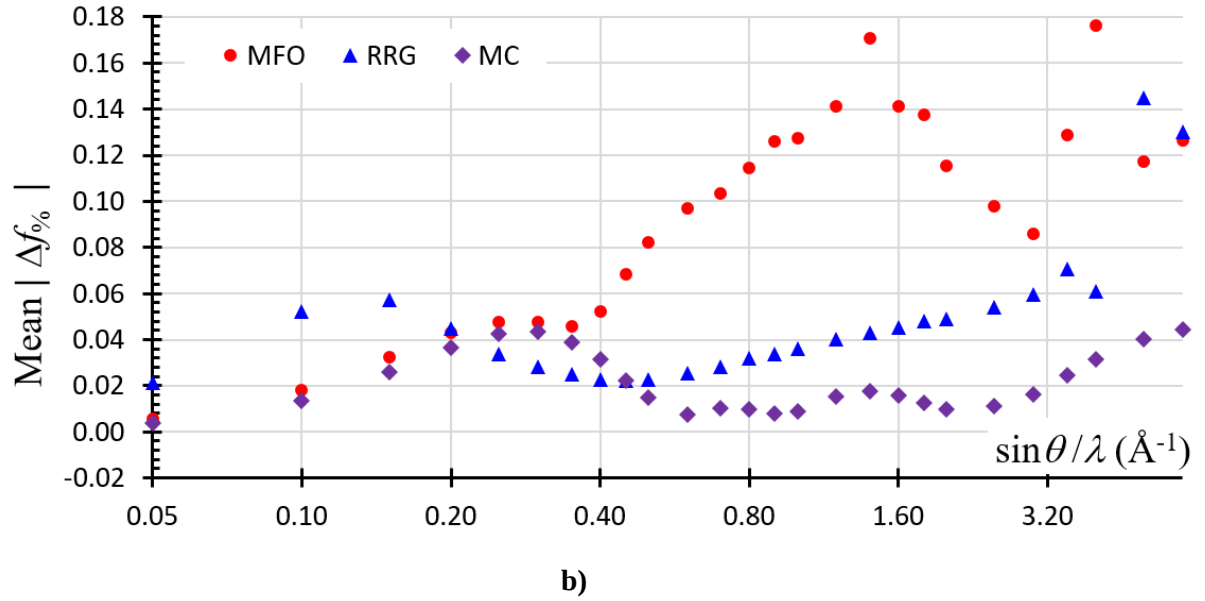


Figure 17 The **(a)** *maximum* and **(b)** *mean differences* in the X-ray scattering factors (Δf) between this work and the previous studies for each $\sin\theta/\lambda$ grid point in the 0 – 6 Å^{-1} range (Rez, Rez and Grant, 1994) for ions common to the studies (Table 5). The previous studies are identified as follows: ● MFO - Maslen, Fox & O'Keefe (2006), ▲ RRG - Rez, Rez & Grant (1994), and ♦ MC - Macchi & Coppens (2001). For convenience, a logarithmic base-2 scale is used for the *x*-axis in both graphs.

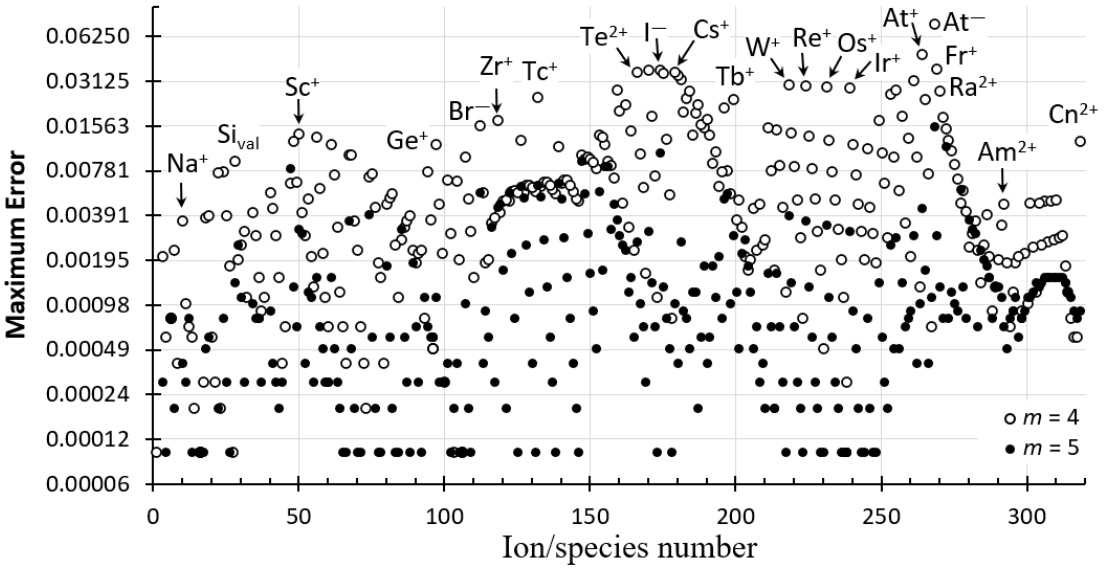


a)

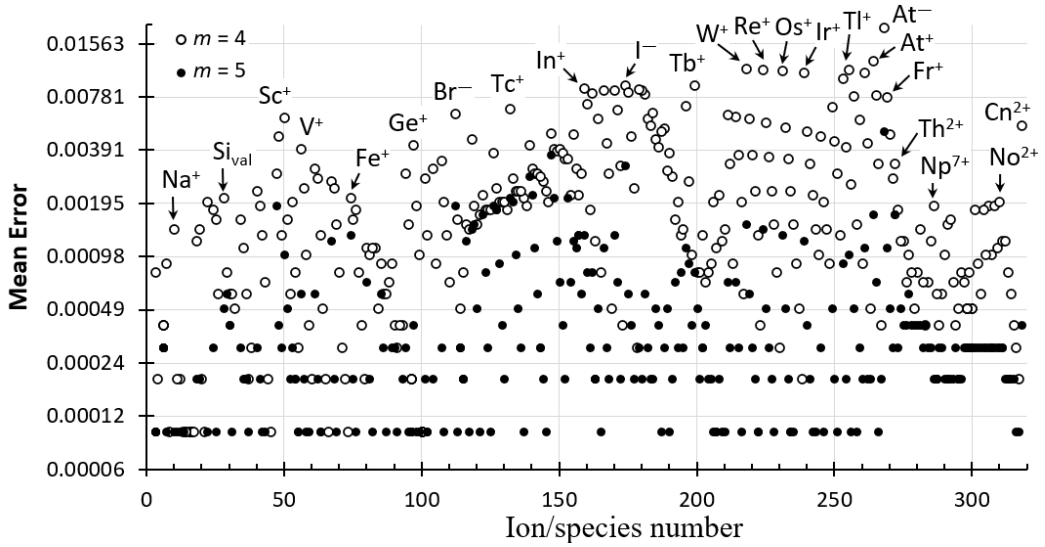


b)

Figure 18 The **(a)** *maximum* and **(b)** *mean percent differences* in the X-ray scattering factors ($\Delta f_{\%}$) between this work and the previous studies for each $\sin\theta/\lambda$ grid point in the 0 – 6 Å^{-1} range (Rez, Rez and Grant, 1994) for ions common to the studies (Table 5). The previous studies are identified as follows: ● MFO - Maslen, Fox & O'Keefe (2006), ▲ RRG - Rez, Rez & Grant (1994), and ♦ MC - Macchi & Coppens (2001). For convenience, a logarithmic base-2 scale is used for the *x*-axis in both graphs.



a)



b)

Figure 19 The differences in the **(a)** maximum and **(b)** mean errors of the interpolating function (50) with $m = 4$ (○) and $m = 5$ (●) (both from this study) for all species included in this work (Table 1). The interpolated $\sin\theta\,/\lambda$ range is $0 - 2$ Å^{-1}. For convenience, a logarithmic base-2 scale is used for the y-axis in both graphs. The zero values (with a precision of four decimal places, 0.0000) are not shown.

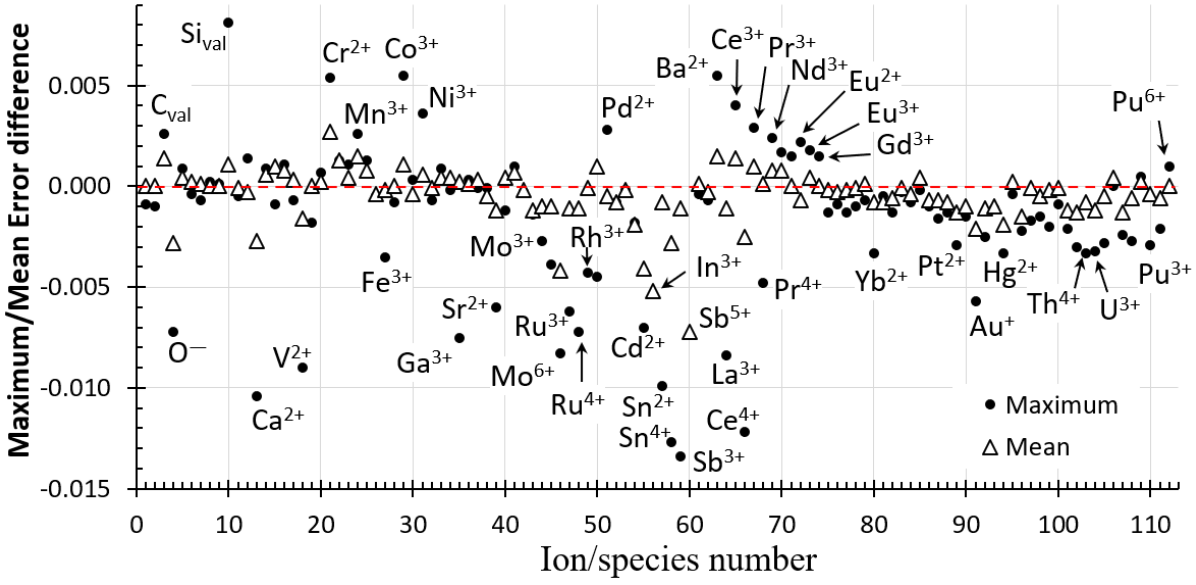
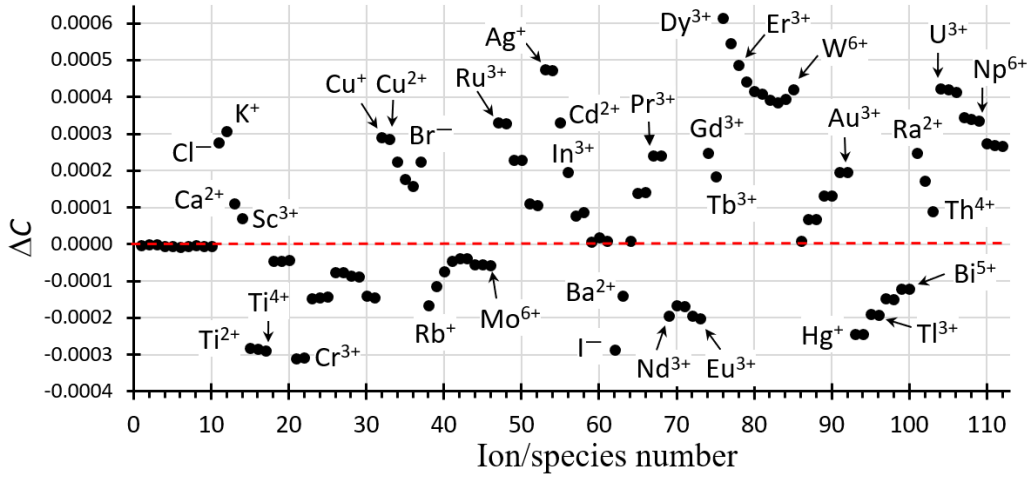
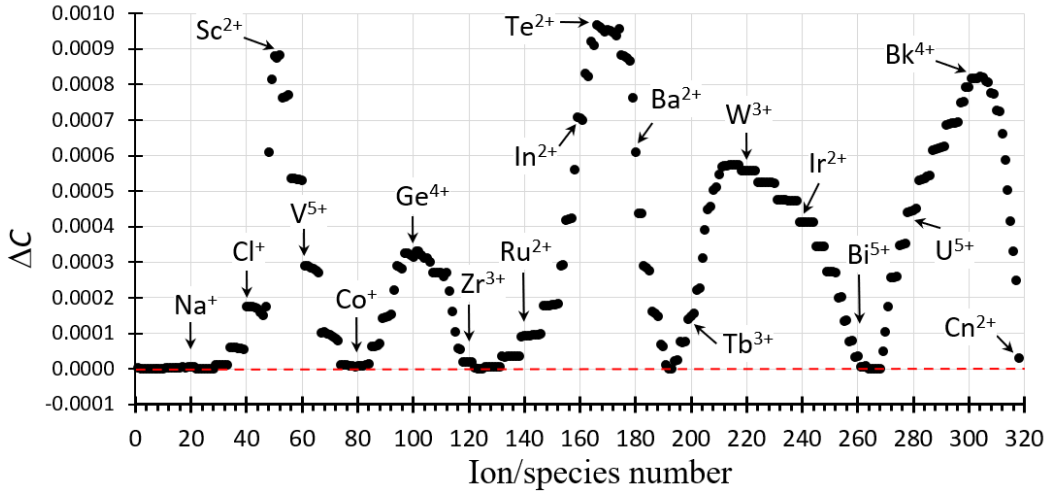


Figure 20 The differences in the maximum (●) and mean (Δ) errors of the interpolating function (50) with $m = 4$ for 112 species (cations, C_{val} and Si_{val}, and anions excluding H^-) relative to the Maslen, Fox & O'Keefe (2006) data. The list of species is given in Table S14. The interpolated $\sin\theta/\lambda$ range is $0 - 2$ Å^{-1}. Negative values indicate a lower error (better fit) in this work. Not shown are the maximum error change of -0.0172 for In^{3+} (number 56) and -0.0277 for Sb^{5+} (number 60).

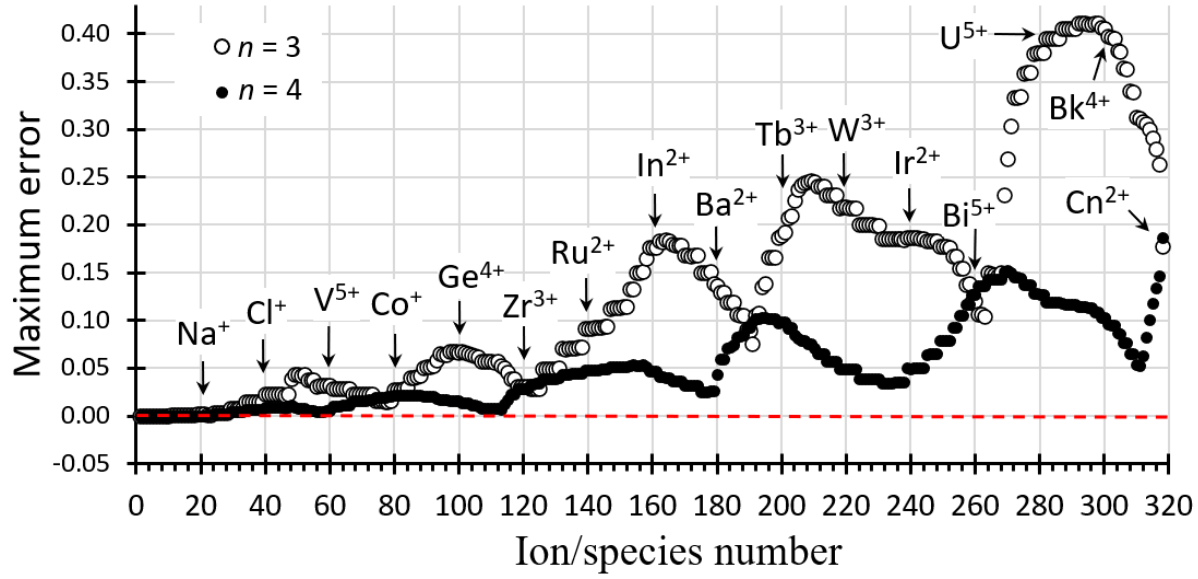


a)

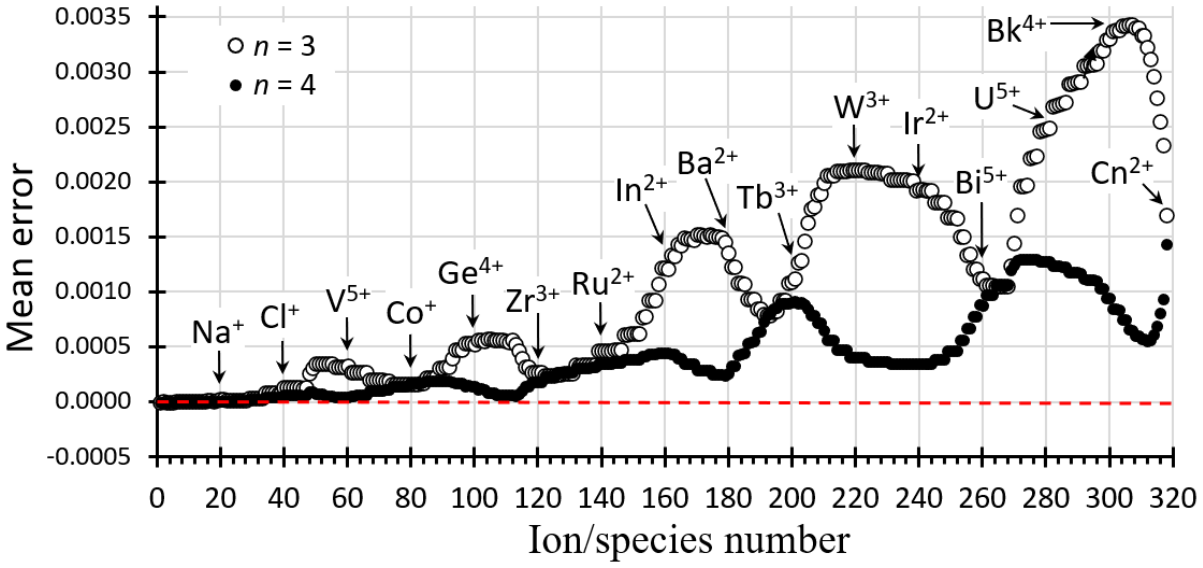


b)

Figure 21 **(a)** The differences in the correlation coefficient C (ΔC) of the interpolating function (51) with $n = 3$ for 112 species (cations, C_{val} and Si_{val}, and anions excluding H^-) relative to the Maslen, Fox & O'Keefe (2006) data. The list of species is given in Table S14. Positive values indicate higher C (better fit) in this work. **(b)** The differences in the correlation coefficient C (ΔC) of the interpolating function (51) with $n = 3$ and $n = 4$ (both from this study) for all 318 species calculated in this work (Table 1). Positive values indicate a higher C (better fit) in the $n = 4$ expansion. The interpolated interval is $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} in both fits.



a)



b)

Figure 22 The **(a)** maximum and **(b)** mean errors of the interpolating function (51) with $n = 3$ (●) and $n = 4$ (○) in the $2 \leq \sin\theta\,/\lambda \leq 6$ Å^{-1} interval for all 318 species calculated in this work (Table 1).

References

Azarov, V. I., Tchang-Brillet, W.-Ü L. & Gayasov, R. R. (2018) *At. Data. Nucl. Tabl.* **121-122**, 345 – 377.

Burden, R. L. & Faires, J. D. (1989). *Numerical Analysis*, 4th ed. Boston: PSW-Kent Publishing Company.

Carlson, T. A., Nestor, C. W. Jr., Wasserman, N. & McDowell, J. D. (1970). *At. Data. Nucl. Tabl.* **2**, 63 –99.

Chiera, N. M., Eichler, R., Steinegger, P., Türler, A., Dressler, R., Piguet, D., Vögele, A., Aksenov, N. V., Albin, Y. V., Bozhikov, G. A., Chepigin, V. I., Dmitriev, S. N., Lebedev, V. Ya., Madumarov, S., Malyshev, O. N., Petrushkin, O. V., Popov, Y. A., Sabel'nikov, A. V., Svirikhin, A. I., Vostokin, G. K. & Yeremin, A. V. (2015). *Towards the selenides of copernicium and flerovium: copernicium - selenium bond formation observed.* In: Türler, A., Schwikowski, M. & Blattmann, A. (eds) *Annual Report 2015, Laboratory of Radiochemistry and Environmental Chemistry*, Paul Scherrer Institut, Switzerland. https://www.psi.ch/sites/default/files/import/luc/AnnualReportsEN/PSI_LCH_AnnualReport2015.pdf

Coulthard, M. A. (1967). *Proc. Phys. Soc.* **91**, 44 – 49.

Cromer, D. T. & Mann, J. B. (1968a). *X-ray scattering factors computed from numerical Hartree-Fock wave functions*. Los Alamos Scientific Laboratory Report LA-3816.

Cromer, D. T. & Mann, J. B. (1968b). *Acta Cryst*. A**24**, 321 – 324.

Cromer, D. T. & Waber, J. T. (1965). *Acta Cryst*. **18**, 104 – 109.

Cromer, D. T. & Waber, J. T. (1968). Unpublished work reported in *International Tables for X-ray Crystallography* (1974), Vol. IV, p. 71. Birmingham: Kynoch Press. (Present distributor: Kluwer Academic Publishers, Dordrecht.)

Desclaux, J. P. (1973). *At. Data. Nucl. Tabl.* **12**, 311 – 406.

Desclaux, J. P., Moser, C. M. & Verhaegen, G. (1971). *J. Phys. B* **4**, 296 – 310.

Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* A**24**, 390 – 397.

Dyall K. G., Grant, I. P., Johnson, C., Parpia, F.A. & Plummer, E. (1989). *Comput. Phys. Comm*. **55**, 425 – 456.

Eichler, R., Brüchle, W., Dressler, R., Düllmann, Ch. E., Eichler, B., Gäggeler, H. W., Gregorich, K. E., Hoffman, D. C., Hübener, S., Jost, D. T., Kirbach, U. W., Laue, C. A., Lavanchy, V. M., Nitsche, H., Patin, J. B., Piguet, D., Schädel, M., Shaughnessy, D. A., Strellis, D. A., Taut, S., Tobler, L., Tsyganov, Y. S., Türler, A., Vahle, A., Wilk, P. A. & Yakushev, A. B. (2000). *Nature* **407**, 63 – 65.

Fong, K. W., Jefferson, T. H., Suyehiro, T. & Walton, L. (1993). *Guide to the SLATEC Common Mathematical Library*. https://www.netlib.org/slatec/guide

Fox, A. G., O'Keefe, M. A. & Tabbernor, M. A. (1989). *Acta Cryst.* A**45**, 786 – 793.

Gäggeler, H. W., Türler, A. (2014). *Gas-Phase Chemistry of Superheavy Elements*. In: Schädel, M. & Shaughnessy, D. (eds) *The Chemistry of Superheavy Elements*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-37466-1_8

Grant, I. P. (1961). *Proc. Roy. Soc. A* **262**, 555 – 576.

Grant, I. P. (1970). *Adv. Phys.* **19**, 747 – 811.

Grant, I. P., Mayers, D. F. & Pyper, N. C. (1976). *J. Phys B: At. Mol. Phys*. **9**, 2777 – 2796.

Grant, I. P., McKenzle, B. J., Norrington, P. H., Mayers, D. F. & Pyper, N. C. (1980). *Comp. Phys. Comm.* **21**, 207 – 231.

Greenwood, N. N. & Earnshaw, A. (1997). Chemistry of the Elements, second edition. Butterworth-Heinemann, Woburn, MA.

Hübener, S., Taut, S., Vahle, A., Dressler, R., Eichler, B., Gäggeler, H. W., Jost, D. T., Piguet, D., Türler, A., Brüchle, W., Jäger, E., Schädel, M., Schimpf, E., Kirbach, U., Trautmann, N. & Yakushev, A. B. (2001). *Radiochim. Acta* **89**, 737 – 741.

Kramida, A., Ralchenko, Yu., Reader, J. & NIST ASD Team (2021). *NIST Atomic Spectra Database* (version 5.9), [Online]. Available: https://physics.nist.gov/asd [Sat Jul 30 2022]. National Institute of Standards and Technology, Gaithersburg, MD. DOI: https://doi.org/10.18434/T4W30F Accessed 30 July 2022.

Johnson, W. R. & Soff, G. (1985). *At. Nucl. Data. Tabl.* **33**, 405 – 446.

Liberman, D., Waber, J. T. & Cromer, D. T. (1965). *Phys. Rev.* **137**, A 27 – A 34.

Macchi, P. & Coppens, P. (2001). *Acta Cryst.* A**57**, 656 – 662.

Mann, J. B. (1968). Los Alamos Scientific Laboratory Report LA-3961.

Maslen, E. N., Fox, A. G. & O'Keefe, M. A. (2006). In *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589.

Olukayode, S., Froese Fischer, C. & Volkov, A. (2022) *Acta Cryst.* A, submitted.

Parpia, F. A., Froese, F. C. & Grant, I. P. (1996). *Comput. Phys. Commun.* **94**, 249 – 271.

Pershina, V., Kratz, J. V., Fricke, B. & Bastug, T. (2000) GSI Scientific Report 1999, 2000-1, 10.

Pershina, V., Kratz, J. V. & Fricke, B. (2000) GSI Scientific Report 1999, 2000-1, 11.

Piessens, R. & de Doncker, E. (1980). *Subroutine QAG in SLATEC Common Mathematical Library*, Version 4.1, July 1993. https://www.netlib.org/slatec/

Redman, S. L., Nave, G. & Sansonetti, C. J. (2014). *Astrophys. J., Suppl. Ser.* **211**, 4-1 – 4-12.

Rez, D., Rez, P. & Grant, I. (1994). *Acta Cryst.* A**50**, 481 – 497.

Rodrigues, G. C., Indelicato, P., Santos, J. P., Patté, P. & Parente, F. (2004). *At. Data. Nucl. Tabl.* **86**, 117 – 233.

Su, Z. & Coppens, P. (1997). *Acta Cryst.* A**53**, 749 – 762.

Su, Z. & Coppens, P. (1998a). *Acta Cryst.* A**54**, 357.

Su, Z. & Coppens, P. (1998b). *Acta Cryst.* A**54**, 646 – 652.

Swirles, B. (1935). *Proc. Roy. Soc. A* **152**, 625 – 649.

Thakkar, A. J. & Smith, V. H. Jr. (1992). *Acta Cryst.* A**48**, 70 – 71.

Vandevender, W. H. & Haskell, K. H. (1982). *SIGNUM Newsletter*, **17**, 16 – 21.

Vand, V., Eiland, P. F. & Pepinsky, R. (1957). *Acta Cryst*. **10**, 303 – 306

Wang, J., Smith Jr,V. H., Bunge, C. F. & Jáiregui, R. (1996). *Acta Cryst*. A**52**, 649 – 658.

Watson, R. E. (1958). *Phys. Rev.* **111**, 1108 – 1110.

Chapter 4. CONCLUSIONS AND FUTURE PLANS[2]

4.1 Conclusions

In the two connected studies presented above, the fully relativistic X-ray scattering factors have been calculated for all neutral atoms with Z = 2 (He) - 118 (Og) and 318 species that included all chemically relevant cations (Greenwood & Earnshaw, 1997), excited (valence) states of the carbon (C_{val}) & silicon (Si_{val}) atoms, and selected monovalent anions (O^-, F^-, Cl^-, Br^-, I^-, At^-), thus significantly extending the coverage relative to all the earlier studies (98 neutral atoms, ≈111 ions, and C_{val} and Si_{val}).

The relativistic one-electron wavefunctions (spinors) were obtained using the B-spline Dirac-Hartree-Fock DBSR_HF code of Zatsarinny & Froese Fischer (2016) while the newly developed Fortran program SF was utilized for high-precision integration of the X-ray scattering factors and determination of the interpolating functions in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} ranges.

Unlike the X-ray scattering factors currently recommended by the International Union of Crystallography (IUCr) and listed in volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) that originate from different sources (Doyle & Turner, 1968; Cromer & Waber, 1968; Cromer & Mann, 1968; Thakkar & Smith, 1992) and were determined at different levels of theory including non-relativistic correlated (Thakkar & Smith, 1992) and Hartree-Fock methods (Cromer & Mann, 1968), and relativistic Dirac-Hartree-Fock (Doyle & Turner, 1968) and Dirac-Slater (Cromer & Waber, 1968) approaches, the newly derived values for all 435 species have been obtained

[2] Several parts of this section also appear in *Acta Crystallographica Section A: Foundations and Advances* (Olukayode, Froese Fischer & Volkov, 2022; 2023)

using the same Dirac-Hartree-Fock level of theory and the same set of approximations that included i) the extended average level (EAL) approach (Grant et al., 1980; Dyall et al., 1989), ii) the Breit interaction correction to the electronic motion due to magnetic and retardation effects, and iii) the Fermi distribution function for the description of a nuclear charge density. While Rez, Rez & Grant (1994) had to resolve to the Watson sphere approximation when calculating anions, the present study did not require that and included a uniform treatment of all cations and monovalent anions.

A detailed comparison of the generated relativistic wavefunctions for neutral atoms in terms of the total and spinor energies, and local and integrated charge density properties with those from a number of previous studies (Rez, Rez & Grant, 1994; Visscher & Dyall, 1997; Guerra et al., 2017; Tatewaki, Yamamoto & Hatano, 2017) confirmed the quality of the calculations. Even though we were unable to compare (due to lack of the available data in literature) the generated wavefunctions for ions the same way it was done for neutral atoms, the total electronic energies were subject to a thorough comparison with the previous theoretical studies (Rodrigues et al., 2004 and to a lesser extent, Macchi & Coppens, 2001) in terms of the absolute values and the atomic ionization energies. While our total energies are much closer to the Macchi & Coppens (2001) values, the determined ionization energies match well those of Rodrigues et al. (2004) and show very similar deviations from the experimentally determined quantities (Kramida et al., 2021).

A B-spline representation of the radial functions combined with a dense radial grid allowed for a precise determination of the X-ray scattering factors values thus avoiding ambiguity associated with the numerical integration procedures used in essentially all the previous investigations. The precision of the integrated X-ray scattering factors was

estimated to include at least eight decimal digits though in the final set of tables we have rounded off the numerical values to five decimals. A comparison of the redetermined X-ray scattering factors with those listed in the 2006 edition of volume C of *International Tables for Crystallography* (Maslen, Fox & O'Keefe, 2006) revealed a number of possible typos and inconsistencies in the published data that have been fixed in our studies.

A thorough comparison of the X-ray scattering factors of ions obtained in this work with those from all the previous studies allowed for a better understanding of the effects of the average / extended average level (AL/EAL) and optimal level (OL) approximations used in the relativistic Dirac-Hartree-Fock calculations. As a reference for future studies, we include in the Supplementary Information a complete list of the employed ground state electronic configurations of the ions with the associated configuration state functions (CSFs) and their weights.

Following the established procedure (Maslen, Fox & O'Keefe, 2006), the X-ray scattering factors in both the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} and the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} ranges have been interpolated using the recommended functions (55)

$$f(\sin\theta/\lambda) = \sum_{i=1}^{m} a_i \exp(-b_i \sin\theta^2/\lambda^2) + c \qquad (55)$$

(Vand, Eiland & Pepinsky, 1957; Doyle & Turner, 1968; Cromer & Waber, 1968; Maslen, Fox & O'Keefe, 2006) and (56)

$$f(\sin\theta/\lambda) = \exp\left(\sum_{i=0}^{n} a_i (\sin\theta/\lambda)^i\right) \qquad (56)$$

(Fox, O'Keefe & Tabbernor, 1989; Maslen, Fox & O'Keefe, 2006) at the conventional ($m = 4$ and $n = 3$, respectively) and extended ($m = 5$ and $n = 4$, respectively) levels. In

comparison, the interpolating functions were not included in the Rez, Rez & Grant (1994) study, while Macchi & Coppens (2001) used the unconventional six-Gaussian expansions in the 0 – 2, 2 – 4, and 4 – 6 Å^{-1} intervals. The extended expansions offer a significant improvement in the accuracy of the interpolated values for the X-ray scattering factors while preserving the general mathematical form of the established equations. For the X-ray crystallographers who require even higher accuracy of the interpolating functions, the SF code can be easily used to extend the function (55) to $m \geq 6$ in the $0 \leq \sin\theta/\lambda \leq 2$ Å^{-1} interval, while for the $2 \leq \sin\theta/\lambda \leq 6$ Å^{-1} range, one can either extend the expansion (56) to $n \geq 5$ or use the function (55) with $m \geq 6$. We also note that the generated relativistic Dirac-Hartree-Fock wavefunctions for neutral atoms and ions stored in the B-spline representations can be easily used to create custom fits of desired accuracy. For example, six-Gaussian fits used in the Su & Coppens (1997, 1998a) and Macchi & Coppens (2001) studies can be readily obtained.

In summary, we believe that the newly derived relativistic Dirac-Hartree-Fock X-ray scattering factors and the accompanied accurate analytical interpolations using the well-established expansions will be useful in the X-ray diffraction studies. For users who require a higher accuracy the extended expansions are presented that require a very minor modification of the existing crystallographic X-ray diffraction software. A detailed analysis of the results suggests that the newly derived values represent an excellent compromise among all the previous studies. The outcomes of the undertaken research should be of interest to the members of crystallographic community who push the boundaries of the accuracy and precision of the X-ray diffraction studies.

4.2 Future Plans

The future plans include

a) calculation of the relativistic Dirac-Hartree-Fock scattering factors for all chemically relevant multivalent anions (Greenwood & Earnshaw, 1997), and

b) determination of the analytical representations of the relativistic Dirac-Hartree-Fock wavefunctions for all neutral atoms and ions using a linear combination of Slater-type functions as was done in the studies by Su & Coppens (1998b) and Macchi & Coppens (2001).

References

Cromer, D. T. & Mann, J. B. (1968a). *X-ray scattering factors computed from numerical Hartree-Fock wave functions*. Los Alamos Scientific Laboratory Report LA-3816.

Cromer, D. T. & Mann, J. B. (1968b). *Acta Cryst*. A**24**, 321 – 324.

Cromer, D. T. & Waber, J. T. (1968). Unpublished work reported in *International Tables for X-ray Crystallography* (1974), Vol. IV, p. 71. Birmingham: Kynoch Press. (Present distributor: Kluwer Academic Publishers, Dordrecht.)

Doyle, P. A. & Turner, P. S. (1968). *Acta Cryst.* A**24**, 390 – 397.

Dyall K. G., Grant, I. P., Johnson, C., Parpia, F.A. & Plummer, E. (1989). *Comput. Phys. Comm*. **55**, 425 – 456.

Fox, A. G., O'Keefe, M. A. & Tabbernor, M. A. (1989). *Acta Cryst.* A**45**, 786 – 793.

Grant, I. P., McKenzle, B. J., Norrington, P. H., Mayers, D. F. & Pyper, N. C. (1980). *Comp. Phys. Comm.* **21**, 207 – 231.

Greenwood, N. N. & Earnshaw, A. (1997). Chemistry of the Elements, second edition. Butterworth-Heinemann, Woburn, MA.

Guerra, M., Amaroa, P., Santos, J. P. & Indelicato, P. (2017). *At. Nucl. Data. Tabl.* **117-118**, 439 – 457.

Kramida, A., Ralchenko, Yu., Reader, J. & NIST ASD Team (2021). *NIST Atomic Spectra Database* (version 5.9), [Online]. Available: https://physics.nist.gov/asd [Sat Jul 30 2022]. National Institute of Standards and Technology, Gaithersburg, MD. DOI: https://doi.org/10.18434/T4W30F Accessed 30 July 2022.

Macchi, P. & Coppens, P. (2001). *Acta Cryst.* A**57**, 656 – 662.

Maslen, E. N., Fox, A. G. & O'Keefe, M. A. (2006). In *International Tables for Crystallography*, vol. C, section 6.1.1, 554 – 589.

Rez, D., Rez, P. & Grant, I. (1994). *Acta Cryst.* A**50** , 481 – 497.

Rodrigues, G. C., Indelicato, P., Santos, J. P., Patté, P. & Parente, F. (2004). *At. Data. Nucl. Tabl.* **86**, 117 – 233.

Su, Z. & Coppens, P. (1997). *Acta Cryst.* A**53**, 749 – 762.

Su, Z. & Coppens, P. (1998a). *Acta Cryst.* A**54**, 357.

Su, Z. & Coppens, P. (1998b). *Acta Cryst.* A**54**, 646 – 652.

Tatewaki, H., Yamamoto, S. & Hatano, Y. (2017). *ACS Omega* **2**, 6072 – 6080.

Thakkar, A. J. & Smith, V. H. Jr. (1992). *Acta Cryst.* A**48**, 70 – 71.

Vand, V., Eiland, P. F. & Pepinsky, R. (1957). *Acta Cryst*. **10**, 303 – 306

Visscher, L. & Dyall, K. G. (1997). *At. Nucl. Data. Tabl.* **67**, 207 – 224.

Zatsarinny, O. & Froese Fischer, C. (2016). *Comput. Phys. Comm.* **202**, 287 – 303. https://github.com/compas/dbsr_hf